EXPLICATION FOR ENGINEERING

Samantha Wakil

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy.

Chapel Hill
2020

Approved by:

Ram Neta

Jack Justus

Alan Nelson

Luc Boven

Jen Morton

ABSTRACT

Samantha Wakil: Explication for Engineering
(Under the direction of Ram Neta)

The conservative idea that it is a philosopher's job to clarify common sense beliefs about ordinary concepts is being weeded out from the population and replaced by a revisionist agenda: philosophers should not merely describe but also analyze and suggest ways to improve our concepts. This project is called "conceptual engineering." The conceptual engineering literature is growing rapidly as more philosophers undertake normative conceptual work. However, many philosophers are practicing conceptual engineering untethered to an explicit methodology. Analyses addressing how we *should* engineer a concept and what *should* be considered successful are scarce. The thesis of this dissertation is that Carnapian Explication solves this methodological deficiency.

For my parents, Gwyn and Bassam Wakil.
They never gave me any idea that I couldn't do whatever I wanted to do or be whoever I wanted to be. But I don't think they ever realized that the person I most wanted to be when I grew up was someone like them. For some reason, my parents always say I'm their inspiration, when in reality they are *the wind beneath my wings*.

ACKNOWLEDGEMENTS

There are numerous ways to describe philosophy and each has a kernel of truth to it. Philosophy is dynamic. It's evolving. It's intellectually challenging. It's inspiring. It is also despairing. Mind-boggling frustrating. And, at times, even demoralizing. I don't consider it an exaggeration to say that philosophy is life-changing. At the very least, it changed my life. Due to the overwhelming amount of love and support I have received from so many people, for six and a half years I have had the honor and privilege to study, and through this dissertation, make my own small contribution to this discipline.

I wish to thank my committee members, professors Ram Neta, Alan Nelson, and Luc Bovens, for two things. First and foremost, for their support and suggestions on this project. And second, for making graduate school fun. Between the many delicious vegan brunches, a tattoo fundraiser, drinks after colloquiums, jazz shows in New Orleans, numerous game nights, and getting to know their incredible families, I have received so much more than intellectual guidance during my graduate studies. I also made lasting memories and friendships with people I hold in such high regard.

It's cliche to say about a person "there are too many things I'm grateful for to mention." But in Jack Justus's case, it's true. So, I'll stick to five, listed in order of increasing importance and gratitude. First, for sharing my love of smoky drinks, hot sauce, and wood chopping. Second, for changing the topic of myundergrad senior seminar from the origin of meaning to the philosophy of math. Third, for having a horrible poker face and making the most bizarre and bewildered expression when I told him I was considering an online master's in education instead of applying to philosophy PhD programs. That alternative arose out of a fear of failing to get in to a good philosophy

department, but that ridiculous face Jack made snapped me out of it. Fourth, for introducing me to Carnap's philosophy. Not only was this dissertation inspired by Carnap's work, but Carnap taught me that "probability is the guide to life," which I consider to be one of the most underrated, but most important, life lessons. Fifth and finally, I thank Jack for providing the kind of research partnership that makes someone believe two minds are really better than one.

Lastly, there are a number of other people who have been instrumental in completing this PhD. Some deserve special mention:

John Roberts—for giving me the harshest feedback I received in graduate school on anything I wrote. My work immensely improved by responding to John's critiques.

Mike Bishop—simply because he is Mike, because he is the best story teller ever, and because it wasn't really 15 minutes.

Simon and Erin—for giving me a home during the Covid-19 pandemic, understanding the need to order take out to avoid cooking and washing dishes, and having a hot tub.

Mike Prinzing—for being my brother in arms in the righteous plight to bridge the philosophy and science divide.

Cheryl—because who could have predicted coming 'out of the closet' and moving into a house together would lead to such a fantastic friendship. I also thank her fur-children, Artes and Sophie. Cheryl would not forgive me if I excluded them.

Jamie—for more Margarita Mondays, lychee martinis, and table-side guacamole orders than I could count. Jamie is also responsible for this dissertation looking like it was written by someone who knows basic spelling and grammar. Finally, I owe Jamie a thank you for letting me hang out with the coolest 5-year-old I have ever met, Naya.

Kaitlin—for exchanging girl power anthems over the years.

Paige—because she's my soul sister, for always dressing up together way fancier than social norms deem appropriate, and for always having each other's backs.

Brandon—whose work ethic always inspired me. And for calling his big sister every once in a while, for advice, which always made me feel like a mature and competent adult. Anyone who has done a PhD knows how rare a feeling that can be.

Ma and Ganoosh—who never went one phone call without saying how proud they are of me, and who will be very relieved when I start a job with a real salary.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Since the topic of this dissertation is Carnapian Explication, a preliminary sketch of the method is required. Most generally, Explication is a procedure for improving ordinary concepts to better achieve specific goals or fulfil particular functions. As Carnap introduces it, "The task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept (or the term used for it) the explicandum, and the exact concept proposed to take the place of the first (or the term proposed for it) the explicatum" (Carnap 1950, 2). Carnap's most detailed description of Explication is in the first chapter of the *Logical Foundations of Probability*. Carnap specifies four adequacy conditions that characterize the procedural component of explication and provide an evaluative schema for gauging success (Carnap 1950, 7):

1. The explicatum is to be *similar* to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts.[1]
3. The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many

[1] Following Carnap, I will use exactness and precision interchangeably.

universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).

4. The explicatum should be as *simple* as possible; this means as simple as the more important requirements (1), (2), and (3) permit.

The goal of this dissertation is to demonstrate how these four criteria provide the methodological foundation that is missing in conceptual engineering. Two objections about the subject matter should be addressed at the outset:

Objection 1: Why do you think conceptual engineering lacks a methodology? Are self-proclaimed engineers like Sally Haslanger, Amie Thomasson, Herman Cappelen, Alexis Burgess, and David Plunkett pursing their projects without recourse to any methodological considerations?

Reply: First, the claim is **not** that there hasn't been *any* methodological analysis in conceptual engineering, but that it is given insufficient attention, thus hindering philosophical progress. Second, conceptual engineers themselves recognize this problem. Here are just three examples:

> If metaphysics centrally involves normative conceptual work, how ought we to be doing it? What methods and standards should we employ? While the idea that work in metaphysics may involve normative conceptual work has begun to gain some traction, far less has been said about how that work is to be done. Answering the methodological question of how it should be done is important. (Thomasson 2020, 436)

> So far I've been talking as though we are looking for a descriptive account of when conceptual revision has gone well . . . We should also be interested in how it *should* be done—i.e., what *should* be considered successful. (Cappelen 2018, 119)

> Obviously some big questions in the vicinity remain entirely unresolved, such as what the proper methodology for conceptual engineering is. (Eklund 2015, 382)

Objection 2: Carnapian Explication is all over the conceptual engineering literature! In fact, many conceptual engineers consider themselves neo-Carnapians. What new contribution can a dissertation on *Explication* make?

Reply: Notice the tension between this objection and the first. Carnap was explicit about Explication being a generalizable methodology: "The task of explication is of very general importance for the construction of concepts. Therefore we shall devote the remainder of this chapter . . . to a discussion of the general nature of the method of explication" (Carnap 1950, 2). The tension should now be clear. Carnap claimed that Explication is a general method for conceptual engineering, and Explication is given considerable homage in the conceptual engineering literature, and yet there are countless complaints that conceptual engineering lacks a methodology. There are two options for resolving this conflict. The first option is to deny the apparent tension. The considerable attention Carnap is given coupled with the recent cry for a methodology suggests that *something* about Explication isn't satisfactory, so conceptual engineers' methodological plight is justified. The second option is to acknowledge that Carnapian Explication is underappreciated and misunderstood by many philosophers today, which has resulted in a failure to recognize the utility of the methodology.

I'm inclined towards option two. Substantial misconceptions about Explication abound in philosophy, even amongst neo-Carnapian conceptual engineers. Amie Thomasson's recent remarks about Explication illustrate this problem:

> Carnap aimed to replace certain everyday concepts with 'exact and fruitful' concepts for use in the formal or empirical sciences . . . While Carnap was interested in conceptual engineering, primarily in the sense of devising new, technical languages, most work in metaphysics (traditional and contemporary) does not involve devising new terms, but rather working with, and making normative choices regarding, common terms of our long-familiar vocabulary. Thus, if we think of metaphysics as engaged in conceptual negotiation regarding terms such as 'freedom', 'person', 'art', 'good', 'responsible', 'number', 'property', 'species', and the like, then we must acknowledge that these are terms that are already part of our shared vocabulary and conceptual scheme—not terms we do or can engineer on a blank slate. So how should we begin? (Thomasson 2020, 44-47)

This quote reveals several points of confusion about Explication. First, Explication is both a procedural method *and* an evaluative schema for the results of the method. Claiming that Explication consists of devising new terms from a "blank slate" without analyzing the familiar terms in our shared

vocabulary completely ignores the procedural component. In the *Logical Foundations of Probability,* Carnap describes how clarification of the original concept (the explicandum) is the first step in undertaking Explications:

> Before we turn to the chief question, viz., what are the requirements for a satisfactory solution of a problem of explication, that is to say, for a satisfactory explicatum, let us look somewhat more at the way in which the problem is to be stated, that is, how the explicandum is to be given . . . we must, in order to prevent the discussion of the problem from becoming entirely futile, do all we can to make at least practically clear what is meant as the explicandum. What X means by a certain term in contexts of a certain kind is at least practically clear to Y if Y is able to predict correctly X's interpretation for most of the simple, ordinary cases of the use of the term in those contexts. It seems to me that, in raising problems of analysis or explication, philosophers very frequently violate this requirement. (Carnap 1950, 4)

Second, while Carnap focused on formal and empirical explications throughout his career, it is a mistake to conclude, on this basis, that Explication is only applicable for those kinds of concepts or that only formal and empirical explications can be successful. The concepts that Thomasson claims interest metaphysicians (species, property, number, freedom, etc.) are all in principle ripe for treatment by Carnapian Explication. In fact, Frege's definition of 'number' is widely taken as an exemplar of explicative methodology (Frege 1953; Carnap 1963b).

The impression that Explication is restricted to formal and empirical concepts seems to be shared by many engineers. Jennifer Nado, for instance, states that "It's worth noting that explication, as Carnap describes it, is primarily tailored towards improvements appropriate to the languages or conceptual schemes of the 'exact' sciences—physics, mathematics, logic, and the like. By contrast, 'conceptual engineering' (at least, as I'll use the term) covers any form of conceptual improvement" (Nado 2019a, 3). Similarly, Herman Cappelen claims that "Carnap's notion of explication, however, is narrower than the activity I'm interested in" (2018, 11). Again, these narrow characterizations are inaccurate. As Carnap specifies in his 1963 reply to Strawson,

> The only essential requirement is that the explicatum be more precise than the explicandum; it is unimportant to which part of the language it belongs. However, since exact concepts are more easily found in the scientific part of our language, it

> will often be useful to define the explicatum in this part . . . The use of symbolic logic and of a constructed language system with explicit syntactical and semantical rules is the most elaborate and most efficient method. For philosophical explications the use of this method is advisable only in special cases, but not generally . . . In my view, the extent to which artificial and elaborate means are used depends on the nature of the philosophical problem in question, and also on the aim of the therapy. The aim may merely be to eliminate an isolated minor difficulty in the simplest way possible. In this case, simple means will suffice. Or the aim may be a more thoroughgoing reform in order to overcome a larger group of interconnected philosophical difficulties. In this case it may be necessary to use more elaborate means and a more comprehensive systematization. (Carnap 1963a, 935-936)

While Carnap did not provide examples of so-called "philosophical explications," the above quote indicates that, nonetheless, he saw Explication as useful for such concepts. This quote was the inspiration for this dissertation. I want to provide what Carnap didn't. This dissertation will explain and demonstrate how the four Explication criteria apply to concepts typically thought by conceptual engineers to fall outside the purview of the methodology, ultimately showing how Explication is the methodology conceptual engineers are searching for.

CHAPTER 2: EXPERIMENTAL EXPLICATIONS

Abstract

This paper argues for two conclusions: (1) evaluating the success of engineered concepts necessarily involves empirical work; and (2) the Carnapian Explication criterion *precision* ought to be a methodological standard in conceptual engineering. These two conclusions provide a new analysis of the race and gender debate between Sally Haslanger and Jennifer Saul. Specifically, the argument identifies the resources Haslanger needs to respond to Saul's main objections. Lastly, I contrast the methodology advocated here with the so-called "method of cases" and draw out some general implications for how we should think about concepts.

2.1 Introduction

The conservative idea that it is a philosopher's job to clarify common sense beliefs about ordinary concepts is being weeded out from the population. In its place, a revisionist agenda is being selected for: philosophers should not merely describe but analyze and suggest ways to improve our concepts. This new research program is called "conceptual engineering" (CE) (see Plunkett and Cappelen, 2020). Conceptual engineering is predicated on two assumptions. First, concepts are the tools we use to understand ourselves and the world around us: "[O]ur conceptual repertoire determines not only what we can think and say but also, as a result, what we can do and who we can be" (Burgess and Plunkett, 2013, 1091). Second, our concepts are not immutable. Many of the concepts we currently possess have changed over time and it's possible for us to investigate whether

our current concepts are defective in some way. Evaluating these concepts, and suggesting ways to improve any defects, is the ultimate goal.[2]

Of course, if we want our concepts to do things and do those things well, then conceptual engineering needs an explicit and effective methodology. This methodological question has, unsurprisingly, become a focal point in recent literature:

> Obviously some big questions in the vicinity remain entirely unresolved, such as what the proper methodology for conceptual engineering is. (Eklund, 2015, 382)

> So far I've been talking as though we are looking for a descriptive account of when conceptual revision has gone well . . . We should also be interested in how it *should* be done — i.e., what *should* be considered successful. (Cappelen, 2018, 119)

> If metaphysics centrally involves normative conceptual work, how ought we to be doing it? What methods and standards should we employ? While the idea that work in metaphysics may involve normative conceptual work has begun to gain some traction, far less has been said about how that work is to be done . . . Answering the methodological question of how it should be done is important.
> (Thomasson, 2020, 436)

Many conceptual engineers have taken Carnap's method of *Explication* as the starting point for methodological inquiry.[3] As Carnap (1950, 2) describes it, "The task of explication consists in

[2] Given the controversy surrounding what concepts are, Cappelen (2018) suggests conceptual engineering should refer to representational devices broadly construed. I don't see any problem with this suggestion, but following the trend in the literature, I will continue to use the word 'concept', keeping in mind that it does not presuppose any views about what concepts are.

[3] There are other varieties of explicative methodology. Carnap's (1950) description was the most developed and thorough at the time (See Carus, 2007). Since then, numerous philosophers have contributed to the method of explication through negative critique (Quine, 1951; Strawson, 1963; Reck, 2012) and friendly-amendments and defense (Brun, 2016; Hanna, 1968; Justus, 2012; Koch, 2019; Maher, 2007; Quine, 1960). A distinguishing feature of Carnapian Explication is that there is no true or correct explicatum: "Carnap's distinction between 'external' and 'internal' questions [has] an obvious application to the process of explication in general. The explicatum . . . belongs to some formalized discourse - some 'framework'. The explicandum . . . belongs ipso facto to a mode of discourse outside that framework" (Stein, 1992, 280). So, for any given explicandum there is no correct explicatum any more than there is a "correct" choice of language. Crucially, though, according to Carnap, we can rationally compare different frameworks. This contrasts sharply with Quine, for whom "there is no stepping outside the 'conceptual scheme' in which our 'mother tongue' places us. We complicate and sophisticate that scheme by means of our science, but we can never be in a position to choose our language" (Carus, 2007, 24). Although other versions of explication may have interesting connections to conceptual engineering, the centrality of Carnap's original conception—and the pluralism at the heart of that conception—aligns much better with the goals of conceptual engineering.

transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept (or the term used for it) the explicandum, and the exact concept proposed to take the place of the first (or the term proposed for it) the explicatum."

Carnap's most detailed description of Explication is located in the first chapter of the *Logical Foundations of Probability*. Carnap specifies four adequacy conditions that characterize the procedural component of explication and provide an evaluative schema for gauging success (Carnap 1950, 7; my emphasis):

1. The explicatum is to be **similar** to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an **exact** form, so as to introduce the explicatum into a well-connected system of scientific concepts.[4]
3. The explicatum is to be a **fruitful** concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).
4. The explicatum should be as **simple** as possible; this means as simple as the more important requirements (1), (2), and (3) permit.

Curiously, despite Carnap's claim that "[t]he task of explication is of very general importance for the construction of concepts" (Carnap 1950, 2), conceptual engineers seem to think Explication does not make good on its promise as a generalizable methodology. Cappelen (2018, 11), for instance, claims that "Carnap's notion of explication, however, is narrower than the activity I'm interested in."

[4] Following Carnap, I will use 'exactness' and 'precision' interchangeably.

Similarly, Pinder (2020, 4) states that "Carnap's proposal is not, as it stands, suitable for this particular task." Assuming these *wholesale* assertions are correct (I will argue in section three that they are not), it's worth considering whether any of the *specifics* from Carnap's proposal could be methodological pillars for conceptual engineering. I argue that experimental philosophy's utility in evaluating conceptual engineering success establishes a demand for the precision criterion in all conceptual engineering projects.

Analyses of Carnapian Explication and experimental philosophy are common features of the conceptual engineering literature. But, despite the prevalence of these two topics, important methodological considerations appear to have been overlooked. The reason is two-fold. First, analyses of Explication have focused almost exclusively on the similarity and fruitfulness criteria, while precision has fallen to the wayside (Lindauer 2019; Koch 2019; Pinder 2017, 2020; Thomasson 2020). Second, arguments about the use of experimental philosophy typically address the 'change of subject' objection (Nado 2019a; Prinzing 2018) or focus on pre-engineering theorizing—similar to what Shepherd and Justus (2015) call "explication preparation" (See also Brun 2016; Nado 2019b; Schupbach 2017).

In section two, I argue that experimental philosophy is not merely useful for the pre-theoretical component of conceptual engineering but also necessary for ultimately evaluating conceptual engineering success. Following a useful distinction made by Koch (2019), this argument applies to theoretical and practical success. In section three, I build off the first argument to establish a second claim: that the third Explication criterion *precision* should be a methodological standard in conceptual engineering. With the exception of Dutilh Novaes (2018), precision has been problematically ignored by conceptual engineers. Section four presents the debate between Haslanger (2000, 2006) and Saul (2006) as a paradigmatic case study of the pitfalls that arise from neglecting precision. I also show how the conclusions established in sections two and three can be

used to respond to Saul's objections of Haslanger's proposed definitions of 'woman' and 'race.' Finally, section five concludes by considering how this analysis relates to the method of cases and responds to anticipated objections.

2.2 The Argument: Part 1

The goal of this section is to argue for the necessity of empirical methods in conceptual engineering. To clarify, I am using the term necessary in an instrumental sense, not a metaphysical or an epistemological sense. This is reminiscent of Quine's famous engineering analogy regarding naturalizing epistemology: "For me, normative epistemology is a branch of engineering . . . The normative here, as elsewhere in engineering, becomes descriptive when the terminal parameter is expressed" (Quine 1986, 664-665). Similarly, the necessity claim in my argument is merely derivative of the stated goals of CE. With this qualification out of the way, here is the argument:

(P1) The goal of conceptual engineering is to evaluate and improve concepts.

(P2) If conceptual improvement is the goal, then the metric for success will be the degree to which an engineered concept makes progress towards achieving its function compared to the original concept.[5]

(P3) Evaluating that success requires information about the functionality of both the original concept and the engineered concept.

(P4) Acquiring that information requires empirical investigation.

(C1) Conceptual engineering necessarily involves empirical work.

[5] Function talk is also rampant in the CE literature. As Nado (2019a) rightly notes, however, the notion of function can be suitably deflated. It need not invoke any commitments to an essentialist notion of telos, or to a telosemantic theory of reference: "it implies no more than the banal fact that we use concepts to do things" (2019a, 11). Accordingly, I will continue to talk about concepts having functions.

Let's proceed in order. Premise one is uncontroversial. A quick glance at the CE literature immediately reveals the goals of the discipline. Here are just three examples:

> Rather than trying to study the concepts we *currently* possess, a philosopher might instead try to determine what concepts we *should* possess. She might attempt to make improvements — to improve clarity or reduce vagueness, to remedy various confusions and inconsistencies in our current concepts, or even to recommend wholesale replacement with a concept that is in some sense superior. (Nado 2019a, 2)
>
> Holding in place some function or functions the concept is to serve, we may aim to redesign old concepts to serve that function better, or to engineer new concepts that can serve a function that was done imperfectly or not at all by our prior conceptual scheme. (Thomasson 2020, 440)

> The terms or concepts which we use to talk and think about a particular subject matter can be defective and can be improved to address these defects . . . If our representational devices can be defective [ ] then we should be engaged in two kinds of activities: (i) investigating whether these concepts are defective and (ii) if defects are found, then ameliorating the defective concepts. (Cappelen 2018, 39-40)

Premise two extends from a straightforward understanding of improvement, which I take to mean something like "making better." Of course, there are numerous ways an engineered concept could be thought of as "better." Perhaps a concept C_x is better than C_y because it's unifying, it's more explanatory, it generates testable predictions, it's simple, etc. Because of this plurality, and because CE is predicated on our concepts *doing* something (i.e., performing some kind of function), I suggest that improvement, more specifically, be understood in terms of 'progress' as defined by Alexander Bird: **Definition of progress**: "If the aim of X is Y, then X makes progress when X achieves Y or promotes the achievement of Y" (Bird 2007, 83). Thus, when Y = [a concept's function] and X = [an engineered concept], X is an improvement over the original concept in so far as it makes progress by achieving or promoting the achievement of Y. Notice this definition of progress is neutral regarding what the function of a concept is. Whether the respective concept's function is the advancement of science (as for Carnap), social and political amelioration (as for Haslanger 2006, 2012), epistemic theorizing (as for Eklund 2002 and Scharp 2013), or explaining moral phenomena (as for Railton 2003), the notion of improvement and thus the general metric for success is the

same: whether and to what degree the engineered concept makes progress (as defined above) over the original.

Premise three is based on the comparative nature of the success standard from premise two. Noticeably, this contrastive element features prominently in Sober's (1999) account of testability. As Sober (1999, 57-58) candidly remarks, when it comes to scientific hypotheses, "Don't ask whether a hypothesis can or cannot explain an outcome. And don't ask whether the hypothesis says that the outcome was probable or improbable. The relevant question is whether the outcome is more probable according to one hypothesis than it is *according to another*. . . We can judge which hypotheses do better and which do worse in their competition, but that is all" (my emphasis). Similarly, we can only judge which concepts do better and which do worse when they're competing. Thus, in order to assess whether an engineered concept is progressive, at least two sources of information are obviously required. First, we need some information about the functionality of the original concept. Second, we need information about the functionality of the engineered concept to make the relevant comparison.[6]

Premise four is the most controversial, especially for those not naturalistically inclined. There are several reasons to accept premise four, starting with another analogy between CE and Philosophy of Science. Besides the claim that hypothesis testing is contrastive, Sober posits that "When biologists test evolutionary hypotheses...The goal is not to determine whether an evolutionary explanation can be invented that is consistent with the theory's delineation of the

[6] Relatedly, this contrastive element also arguably undergirds Carnap's caution against conflating explicanda with explicata. As his diagnosis of the problem of probability indicates, "When we look at the formulations which the authors themselves offer in order to make clear which meanings of 'probability' they intend to take as their explicanda, we find phrases as different as 'degree of belief,' 'degree of reasonable expectation,' 'degree of possibility,' 'degree of proximity to certainty,' 'degree of partial truth,' 'relative frequency,' and many others. This multiplicity of phrases shows that any assumption of a unique explicandum common to all authors is untenable . . . But most investigators in the field of probability apparently believe that all the various theories of probability are intended to solve the same problem and hence that any two theories which differ fundamentally from one another are incompatible. . . These mutual rejections are often formulated in rather strong terms. This whole controversy seems to me futile and unnecessary. The two sides start from different explicanda" (Carnap, 1945, 517-518).

possible causes of evolution, but to determine whether there is evidence that discriminates between different evolutionary hypotheses concerning what actually occurred" (Sober 1990, 60). Perhaps surprisingly, evaluating engineered concepts appears strikingly similar to testing evolutionary hypotheses. The goal is not to devise new terms from a "blank slate" that are merely consistent with a concept's possible function (Thomasson 2020, 447). Rather, the goal is to acquire evidence that discriminates between different explicata regarding which *actually* achieves or promotes the achievement of the concept's function. The best source of that evidence is found through experimental philosophy.

Experimental philosophy uses some of the empirical methods of social psychology and cognitive science to acquire data relevant to addressing traditional philosophical questions and debates (See Plakias 2015 for a useful overview). It is typically divided into two programs. The negative program appeals to empirical results to cast deep skepticism about intuitions being the primary source of philosophical evidence. Although the debate about their epistemic significance has a long and contentious history, a new cluster of worries about the reliance on intuition arguably began with the experimental studies by Weinberg et al. (2001), Knobe (2003), and Alexander and Weinberg (2007). For example, the so-called "Knobe effect" revealed an asymmetry in participants' attribution of intentional action. (More on this below.) Additionally, there are disparate intuitions among philosophers regarding scenarios involving concepts of interest like free will, knowledge ascriptions, moral blame, twin earth cases, etc. Importantly, it appears this same kind of diversity in intuitions is found among non-philosophers (Weinberg et al., 2001).

Philosophers in the negative program also argue that intuitions are sensitive to philosophically irrelevant factors like a person's gender and cultural and socioeconomic background, the way an outcome is described, or the order in which the cases are presented. For example, Swain et al. (2008) found that attributions of knowledge were subject to order effects. When participants

were presented a clear case of knowledge before a true temp case, they were less likely to attribute knowledge to true temp than participants who received the true temp case first. Yet, when participants received a clear non-knowledge case first, they were more likely to attribute knowledge to the true temp case than participants who received true temp first. In sum, experimental philosophy's negative program has made an aggressive assault on intuition-based theorizing. The experimental results from the studies just cited (and many others) are taken as demonstrating the unreliable, relative, and highly sensitive nature of intuitions, making them a poor evidential source for philosophical theorizing and thus warranting a significant change to our philosophical practices.[7]

Against the naysayers, the positive program uses experimental data to gain information about folk concepts and intuitions in order to *support* traditional answers to philosophical questions. Philosophers in the positive program are "motivated to explore intuitions experimentally because they think that by doing so they can do a better job of conceptual analysis" (Stich 2016). For example, Neta (2012, 333) argues there is an intuitively compelling rationale that predicts the order effect found in Swain et al. (2008): "It is intuitive that, for a psychologically salient property X (e.g., how substantial a defeater is), when a borderline case of X is compared against clear cases of X, it seems less X-ish, and when a borderline case of X is compared against clear cases of non-X, it seems more X-ish." Of course, one example does not vindicate the reliability of intuition. The claim that philosophers can, in general, reliably predict what ordinary subjects will and will not find intuitive is itself an empirical question, and one the positive program can address.

Furthermore, some recent studies have not replicated the original findings of the negative program (Seyedsayamdost 2015; Adleberg et al. 2015) and other studies have found that there is some cross-cultural uniformity in knowledge attributions in Gettier cases (Machery et al. 2015).

[7] See also Cola¸co et al., 2014; Cruz, 2017; Machery et al., 2004; May et al., 2010; Nichols and Knobe, 2007; Schwitzgebel and Cushman, 2015; Sinnott-Armstrong, 2008; Tobia, Buckwalter, and Stich, 2013.

Moreover, the positive program emphasizes how experimental philosophy can uncover information about our concepts that was previously unconsidered by philosophers. As Plakias (2015, 8; my emphasis) summarizes, for example,

> Strohminger and Nichols (2014) presented subjects with the hypothetical case of someone named Jim, who had been in a car crash and suffered psychological damage. They then described various changes to the person's personality or memory and asked subjects whether he remained the same person or whether his identity had changed. What they found was that a change in Jim's moral attitude was the most influential factor in determining that his identity had changed—more influential than memories, desires, or other personality traits. *This finding raises the novel possibility that the ordinary concept of the self includes a moral dimension*, such that one's moral views are an essential part of one's personal identity.

Despite their differences, taken together, the negative and positive programs demonstrate how "[f]or every philosophical use of intuition, there is a role for experimental philosophy" (Plakias, 2015, 7). The experimental methods don't merely provide information about our concepts; crucially, they "reveal the contours of concepts" (ibid., 6). Thus, for the purposes of CE the distinction between the negative and positive programs appear obsolete. Whether the results of a study are taken to confirm or disconfirm the reliability of a particular intuition, the crucial point remains: experimental philosophy provides a testable, scientifically-grounded, and data-driven means for acquiring the evidence needed to evaluate the functionality of a concept.

Let me address two objections before moving on. First, one might reject the conclusion of this argument on the grounds that CE is best construed as a form of ideal-theorizing. Meta-semantic considerations, and semantic externalism in particular, have raised concerns about actually implementing suggested conceptual revisions. If one thinks, as the semantic externalist does, that people have very little control over the meanings of our terms, then the feasibility of changing a concept according to an engineer's recommendation will seem highly improbable, if not impossible (Plunkett and Cappelen 2020). The concern about feasibility has motivated some to suggest that CE should not be concerned with implementation at all. Perhaps, as Burgess and Plunkett (2013, 1096)

suggest, "Whether we ought to use a given concept could be completely independent of how hard or easy it would be to do so." Cappelen similarly argues that conceptual engineers should be more like political theorists who "are comfortable reflecting on and proposing theories of justice, for example, without having a recipe for how they can be implemented. More generally, we can make judgments about what ought to be the case, without knowing how to make the world that way (or even having a plausible strategy in mind). Those are just different projects: figuring out what ought to be the case and making the world that way" (Cappelen 2018, 83-4). This objection misconstrues the scope of my argument. The argument is not meant to convince semantic externalists that conceptual revision is possible (nor does it assume that it is!). The argument should be viewed conditionally. If the goal is conceptual improvement, then the argument specifies what is required for evaluating said improvement. If one doesn't care about improvement, or if meaning change really is untenable, the conclusion will seem inconsequential. That said, semantic externalists should welcome this kind of argument even with their doubts about feasibility. The empirical work needed to assess the success of CE projects will have implications for confirming or disconfirming semantic externalism. If the majority of CE projects are unsuccessful, this could be evidence the semantic externalists point to in arguing that CE should be done as a form of ideal-theorizing. Analogously, CE as non-ideal theorizing could be vindicated by data showing that the recommended conceptual changes are in some cases actually successful. I have no hunches about which way it will go; we have to let the data speak. In order for that to happen, conceptual engineers need to take steps in acquiring the relevant data.

Another objection might be made in light of a distinction between practical success and theoretical success (Koch 2019). For Koch, theoretical success is conditional: "if people were to use it [the engineered concept], they would make epistemic progress" (ibid., 709). Koch also seems to support the purely normative inquiry that Burgess, Plunkett, and Cappelen advocate for: "In general,

which concepts we do use is often influenced by factors which have nothing to do with whether we should use them" (ibid., 709). In contrast, practical success concerns "whether people are actually ready to use the concepts they should use" (ibid., 709). Koch departs from Burgess, Plunkett, and Cappelen in thinking the normative question is the only one relevant for CE: "If one's goal is to make actual progress in a given domain, both success conditions matter" (ibid., 709). However, Koch thinks the success conditions for the theoretical and the practical are different and should not be conflated:

> Even if you believe that in order to do conceptual engineering well, you need to pay attention to both practical and theoretical standards, you are well-advise to keep them separate. This is because the kind of factors which determine whether a concept meets the theoretical success condition is usually very different from the kind of factors which determine whether it meets the practical one. What matters for theoretical success are broadly epistemic factors: whether the concept in question allows us to make useful and novel distinctions, whether it allows us to secure the progress we already have made, whether it allows us to view a matter in a new and interesting light, etc. What matters for practical success, on the other hand, are psychological and sociological factors: whether the concept triggers positive reactions, whether it 'sticks' to people's minds, whether the people advocating it has enough influence or political power to make it spread, etc. Given that these factors are typically very different, investigations into them take different forms. Evaluating theoretical success is often an issue for philosophers of science or for the philosophically-minded subgroup of the theoretical community in which the concept is used. Evaluating practical success, on the other hand, will typically be a matter of social psychology and sociology. (Koch 2019, 709-710)

With this distinction laid out, a potential objection to the argument is that its scope is limited to practical success. One might agree with Koch that the empirical work done in social science is for evaluating practical success rather than theoretical success, which is not determined experimentally. Even if my argument was limited to practical success, that is still considerable progress. Both kinds of success matter, so an argument that elucidates one of the two components for successful CE projects is a valuable contribution. Nonetheless, the claim that empirical methods are not germane to evaluating theoretical success is dubious. Whether or not a concept has certain theoretical virtues is in many cases (though perhaps not all) an empirical question.

Take, for example, the presumed theoretical virtue of parsimony. Parsimony can be defined in a variety of ways against a variety of objectives. There are at least three in the philosophy of science literature: a quantitative concept that refers only to the number of processes, entities, causes, or basic principles postulated; a qualitative concept that refers to the number of types of processes, entities, causes, or basic principles postulated; and a Bayesian concept based on Reichenbach's principle of common causes (Sober, 2015). Whether a particular explanation or theory is *simple* according to any of these definitions is not a matter of mere philosophical reflection. Each of these concepts gives very specific criteria by which to gauge simplicity and such evaluations require "looking out" to determine if those criteria are met.[8] Thus, it's not surprising that parsimony claims in evolutionary biology are often challenged or overturned in light of new phylogenetic analyses, ecological data, and other relevant experimental findings.[9]

For another example, consider the Knobe-effect mentioned before. Knobe (2003) asked participants to consider a scenario where the boss of a company decides to start a new program to increase profits knowing there will be certain foreseen side effects in implementing the program. Depending on whether the known side effect either helped or harmed the environment, participants gave opposing answers to the question of whether the boss *intentionally* brought about the side effect. Recall Koch's first condition for theoretical success was that "the concept in question allows us to make useful and novel distinctions" (2019, 709). One thing Knobe's **experimental** results did was to reveal a previously unknown feature about the concept of intentionality that was central in the

[8] Machery (2017) makes a similar point related to what he calls "prescriptive conceptual analysis." Machery argues that given some specified set for normative constraints, experimental philosophy can help determine which concept proposals meet those constraints.

[9] A great illustration of the connection between empirical work and concept determination in ecology can be found in Justus, 2012.

philosophy of action and ethics. Before Knobe's studies, the prevailing view was that an agent must have intended to perform an action in order for that person to be held morally responsible. What the data surprisingly revealed, however, was that the moral valence of a consequence influenced participants' intentionality judgments. Most participants believe an act is intentional if the consequence is harmful, and most believe an act is *not* intentional if the consequence is helpful.

Relatedly, Koch's second criteria for theoretical success, "whether it allows us to secure the progress we have already made," is also amenable to empirical investigation. It is important to ask exactly how much theoretical progress we have made and to gauge that progress with the relevant data. Sober explains this point nicely with regards to the theoretical virtue unification: "We must avoid the mistake of selective attention; we need to count the failures as well as the successes . . . we need to watch out for a selection bias; the apparent track record to date overestimates the success rate of unification and it is hard to say by how much" (Sober 2015, 46-47). Again, Knobe's study compels us to question the presumed progressiveness of the 'intentionality' concept in the philosophy of action. Moreover, ameliorative or revisionary CE projects start from the presupposition that our current concepts are defective insofar as they have impeded certain forms of progress. Empirical work is used to support such an assumption. For example, empirical research on gender and racial bias reinforce this starting point for Haslanger's (2000) analysis and suggested revisions of gender and race concepts. So, I want to maintain the importance of Koch's distinction while recognizing that evaluating theoretical progress, and determining whether our engineered concepts have contributed to that progress, can (and should) be informed by empirical investigation.

2.3 The Argument: Part 2

The argument in the previous section sought to establish the necessary role of empirical methods for evaluating the success of CE projects. The conclusion that CE necessarily involves empirical work suggests a related methodological consideration. As mentioned in the introduction,

Carnapian Explication is given extensive homage in the CE literature. However, Explication is typically characterized as a subtopic of CE with its application being wrongly perceived as limited to scientific and logical concepts:

> Carnap aimed to replace certain everyday concepts with 'exact and fruitful' concepts for use in the formal or empirical sciences . . . Carnap was interested in conceptual engineering, primarily in the sense of devising new, technical languages. (Thomasson 2020, 6)

> Carnap's notion of explication, however, is narrower than the activity I'm interested in. (Cappelen 2018, 11)

> For some purposes, formulating either empirical laws or logical theorems is not what really matters - just think of typical philosophical inquiries, e.g. into the nature of justice or knowledge. (Koch 2019, 702)

> Carnap's proposal is not, as it stands, suitable for this particular task. (Pinder 2020, 4)

> It's worth noting that explication, as Carnap describes it, is primarily tailored towards improvements appropriate to the languages or conceptual schemes of the 'exact' sciences — physics, mathematics, logic, and the like. By contrast, 'conceptual engineering' (at least, as I'll use the term) covers any form of conceptual improvement. (Nado 2019a, 3)

Although Carnap focused on logical and empirical explications throughout his career, it's a mistake to conclude that Explication, as Carnap described, is only applicable for those concepts or that formal explications could only be deemed successful. Carnap did not see Explication as carving out a joint between philosophical and scientific concepts: "I see here no sharp boundary line but a continuous transition" (Carnap 1963, 934). Further, he explicitly acknowledged the limits of formal methods for such purposes. This is clear from his reply to Strawson:

> The use of symbolic logic and of a constructed language system with explicit syntactical and semantical rules is the most elaborate and most efficient method. For philosophical explications the use of this method is advisable only in special cases, but not generally . . . Again misled by his sharp distinction between scientific language and ordinary language, Strawson seems to misunderstand this point. He believes that the explicatum is meant to serve a scientific purpose, in distinction to the explicandum which serves a pre-scientific purpose. (Carnap 1963, 935)

Recall that premise two of the first argument was neutral with respect to the function of our target concepts. Thus, the scope of the argument extends across all varieties of CE projects (scientific progress, political amelioration, etc.). This wide range coupled with the textual evidence that Carnap saw Explication as a generalizable methodology motivated my suspicion that conceptual engineers have missed something important. In particular, discussion of the third Explication criterion *precision* is strikingly absent from the CE literature. At one point Carnap even claims that "[t]he only essential requirement is that the explicatum be more precise than the explicandum" (1963, 936). Perhaps conceptual engineers have mistakenly assumed that precision was included for its importance in explicating scientific and logical concepts, and thus wouldn't be relevant for non-scientific CE tasks. Understanding why Carnap included precision as an Explication criterion reveals why precision is pertinent to all CE projects.

Carnap's claim that precision is the "only essential requirement" seems to suggest that he viewed it as the most important condition for an Explication. However, several scholars have made clear that Carnap saw fruitfulness as the ultimate success metric and thus gave it the most evaluative weight (see Brun 2016; Carus 2007; Justus 2012; Dutilh-Novaes 2018; Schupbach 2017). Precision is included and given significant attention because of its relationship to fruitfulness. As Justus (2012, 168) explains, "precision is paramount because it *usually* enhances fruitfulness." Although Carnap denied a sharp philosophy and science divide, he took the unparalleled epistemic success of science as a model to inform philosophical methodology. Carnap's great insight was that vague concepts are rarely components of well-confirmed generalizations. Increasing precision, on the other hand,

> often facilitates in developing and discovering well-confirmed generalizations. . . Without sufficiently precise concepts, it is difficult if not impossible to derive predictions from statements containing them. Without such predictions, in turn, statements cannot be confirmed or disconfirmed. Testing predictions of hypotheses and theories against data is a staple of scientific methodology that has clearly proved to be an epistemically reliable basis for inference. Precision facilitates this methodology and its success therefore grounds the exactness [precision] criterion. (Justus 2012, 169)

How exactly does precision reliably increase fruitfulness? The history of scientific progress suggests that increasing precision "usually enhances mathematical rigor, measurability, testability, [and] theoretical unification..." (Shepherd and Justus 2015, 338). Although this list is targeted at empirical concepts, there are two items that are crucial for any conceptual engineering project—measurability and testability. As premise two of the previous argument made clear, if we want to evaluate conceptual improvement, then regardless of the kind of target concept (scientific, social, political, metaphysical, etc.), we must have some way to measure/test the degree to which the engineered concept functions better or worse than the original.[10] Thus, the likelihood that more precise concepts will increase measurability and testability provides a strong motivation for making precision an essential component of CE methodology. Here is the whole argument:

(P1) The goal of conceptual engineering is to evaluate and improve concepts.

(P2) If conceptual improvement is the goal, then the metric for success will be the degree to which an engineered concept makes progress towards achieving its function compared to the original concept.

(P3) Evaluating that success requires information about the functionality of both the original concept and the engineered concept.

(P4) Acquiring that information requires empirical investigation.

(C1) Conceptual engineering necessarily involves empirical work.

[10] Notably, in his work on Interpretation and Preciseness (1950), Arne Naess endorses the importance of precision in testing and modifying conceptual frameworks. Naess also undertakes his own experimental philosophy studies on folk uses of various semantic concepts like 'synonymity.' The results of these studies lead Naess to a similar conclusion of the first argument presented here: "acceptance of intuitions reported by the philosophically sophisticated about the verbal and conceptual habits of others leads to confusion and error. . . empirical procedures should be applied to empirical questions. When philosophers offer conflicting answers to questions that have empirical components, empirical research is needed." Thanks are owed to an anonymous reviewer for bringing Naess's work and Carnap's (1955) praise of that work to my attention.

(P5) Given the empirical necessity for evaluating the success of engineered concepts, we ought to engineer concepts that are empirically measurable/testable.

(P6) Increasing precision *typically* facilitates the evaluation of a concept's functionality by increasing experimental measurability/testability.

(C2) Precision ought to be a methodological standard in conceptual engineering.

The rationale behind premise five is straightforward: if we care about evaluating the success of engineered concepts, then we should engineer concepts with features that aid in that evaluation. Again, if you don't think feasibility really matters for CE, then you won't find this compelling. But most conceptual engineers do care about actual success, especially those interested in scientific and sociopolitical progress. The justification for premise six was given above by clarifying Carnap's rationale for including precision as an Explication criterion. Two clarifications about the argument are noteworthy.

First, one might think my argument does not provide any reason to think that it's easier to assess function-achievement of precise concepts than imprecise ones. What the argument shows is that conceptual engineers need to be more precise about 'concept function' or 'function achievement' in order to investigate concept functionality. While the notions of concept function or concept achievement are, of course, amenable to more precification, without the *concept* itself being sufficiently precise it will be difficult to formulate predictions from statements, descriptions, hypothesis, or explanations containing them. Without such predictions, the concept's functionality cannot be confirmed or disconfirmed.

Second, notice the preceding remark said *sufficiently* precise. This is important because it would be a mistake to think the argument requires all concepts to be maximally precise. Carnap characterized non-logical concepts in order of increasing precision: classificatory, comparative, and quantitative (Carnap 1950, 12). Even though increasing precision along these dimensions will

typically facilitate evaluating concept functionality, it is a truism that this will not always be the case. The argument allows for decreases of precision when warranted. Again, for Carnap, "Precision for precision's sake is not the agenda. Rather, enhancing precision usually enhances fruitfulness, which is the agenda" (Shepherd and Justus 2015, 388). In cases where the added precision does not result in increases in fruitfulness, or even decreases fruitfulness, then the focus should be on finding fruitful comparative or classificatory explicata (Carnap 1950, 14). Nonetheless, it remains necessary for accessing conceptual improvement that we can measure functionality. In cases where quantified methods seem ill-suited, or are unavailable, we can use qualitative measures. Given that some concepts of interest to conceptual engineers might not be improved by quantification, this flexibility is a feature and not a bug of the argument. To buttress this conclusion, the next section will examine a well-known debate in the literature that I argue demonstrates how overlooking precision impedes progress in CE.

2.4 Carnap vs Strawson, *again*.

The title of this section might suggest that I will discuss the 'change of subject' objection Strawson (1963) famously levels against Carnap, but I want to analyze the modern rendition of this dispute that occurs in the feminist literature on race and gender. In this production, Jennifer Saul and Sally Haslanger are our contemporary equivalents of Peter Strawson and Rudolf Carnap, as illustrated by a brief comparison of the two parties. Strawson and Saul are conservatives; in Carnapian terms they prioritize similarity:

> The actual use of linguistic expressions remains his [a philosopher's] sole and essential point of contact with reality; for this is the only point from which the actual mode of operation of concepts can be observed. (Strawson 1963, 508)

> It seems to me that communication is difficult enough as it is, and that we should instead try to use ordinary terms in as ordinary a way as possible. (Saul 2006, 141)

Relatedly, Strawson and Saul also seem critical of formal or technical explications in general:

> To offer formal explanations of key terms of scientific theories to one who seeks philosophical illumination of essential concepts of nonscientific discourse, is to do something utterly irrelevant — is a sheer misunderstanding, like offering a text-book on physiology to someone who says (with a sigh) that he wished he understood the workings of the human heart. (Strawson 1963, 505)

> Haslanger's definitions are, as she concedes, at odds with intuitions. For most of us, Haslanger's definitions are just not the sort of thing we are likely to come up with if asked about the nature of race and gender, and when presented with them we will very likely reject them. (Saul 2006, 122)

> We may well have a need for technical terms that work as Haslanger's do, but it seems to me a mistake to risk confusion by using ordinary vocabulary to do this work. (Saul 2006, 141)

Carnap and Haslanger, on the other hand, are revisionaries. What they value most is fruitfulness:

> When we compare the explicandum Fish with the explicatum Piscis, we see that they do not even approximately coincide . . . The change which zoologists brought about in this point was not a correction in the field of factual knowledge but a change in the rules of the language . . . they realized the fact that the concept Piscis promised to be much more fruitful than any concept more similar to Fish. (Carnap 1950, 6)

> I've cast my inquiry as an analytical — or what I here call an ameliorative — project that seeks to identify what legitimate purposes we might have (if any) in categorizing people on the basis of race or gender, and to develop concepts that would help us achieve these ends. I believe that we should adopt a constructionist account not because it provides an analysis of our ordinary discourse, but because it offers numerous political and theoretical advantages. (Haslanger 2012, 366)

Although Haslanger and Carnap obviously have different subject areas in mind (politics for Haslanger and science for Carnap), they clearly share a commitment to improving our concepts to achieve various goals. I am not going to focus on the change of subject objection or, more generally, whether similarity or fruitfulness should be given the most weight. There are already plenty of arguments in the literature regarding that particular trade-off. What is interesting about the debate between Saul and Haslanger is that a close examination of the dialectic reveals why precision is paramount for CE.

First, a very general remark: if we invoke only two criteria for assessing the success of an engineered concept, then any time the two criteria diverge from each other we are left at a standstill.

The Saul-Haslanger and Strawson-Carnap debates are presented as involving a false dichotomy: we either go with a fruitful explicata *or* go with a similar explicata. But if we have more than two criteria, then those other standards will be brought to bear on our conceptual decision-making. For example, if one explicata ranks high in fruitfulness and another high in similarity, we should also look at how they measure against simplicity and precision. Of course, using four explication criteria as opposed to two does not guarantee there will never be conflict; simplicity and precision can often pull apart the way fruitfulness and similarity do. Trade-offs between the four criteria will always have to be made, and on a case-by-case basis. The point is simply that rather than making our conceptual decisions based on a false dichotomy between fruitfulness and similarity, remembering to include all four explication criteria gives us much more evaluative capability.[11]

Second, aside from the problem of a false dichotomy, it appears that failing to account for precision has left Haslanger's analysis of gender and race open to a set of objections concerning measurability.[12] Saul (2006, 138), for example, claims that when it comes to the supposed benefits of Haslanger's proposed definitions, "there are many uncertain empirical matters." Curiously, it seems that lack of empirical tractability is one of Saul's main points of contention:

> How, then, could we ever hope to study our actual usage of race concepts, as opposed to colour concepts? . . . It strikes me as unlikely that we would have much evidence enabling us to attribute such a distinction in a non-question-begging manner, collecting together one set of usages of 'Black' as telling us about a race concept and another as telling us about a colour concept. (ibid., 132)

[11] See Olsson, 2015 for a compelling case study using the four explication criteria to adjudicate theories in epistemology.

[12] Haslanger's proposed definition of 'woman' is the following: "S is a woman if S is systematically subordinated along some dimension (economic, political, legal, social, etc.) and S is 'marked' as a target for this treatment by observed or imagined bodily features presumed to be evidence of a female's biological role in reproduction" (2000, 39). Similarly, for race, "A group is racialized if its members are socially positioned as subordinate or privileged along some dimension (economic, political, legal, social, etc.) and the group is 'marked' as a target for this treatment by observed or imagined bodily features presumed to be evidence of ancestral links to a certain geographical region" (ibid., 44).

> In order to support this view, we would need to find behavioural evidence that speakers do in fact classify people in this way. It seems to me that such evidence is unlikely to be forthcoming, given both the muddled linguistic usage and its tendency to go unnoticed. (ibid., 131)
>
> To learn about the operative concept, what we really need to do is find some females who are not subordinated on the basis of real or perceived sex characteristics and check to see whether people apply the term 'woman' to them. But this is far from straightforward. . . (ibid., 129)
>
> Haslanger recognizes that the proponent of a revisionary analysis faces an additional issue: when does an analysis cross the boundary from being revisionary to being an analysis of something entirely different? . . . [S]he suggests two criteria for deciding this, a semantic and a political one . . . The semantic condition, she suggests, is far more straightforward: 'the proposed shift in meaning of the term would seem semantically warranted if central functions of the term remain the same, e.g., if it helps to organize or explain a core set of phenomena that the ordinary terms are used to identify or describe.' Haslanger seems to think that her analyses of race and gender terms will pass these tests, although she does not argue this in detail . . . *it is far from obvious how the phenomenon in question should be specified.* (ibid., 134; my emphasis)

It is crucial to situate these criticisms in the broader context of the Saul-Haslanger debate in order to appreciate their bite. Saul's critique here is primarily referring to what Haslanger calls the "descriptive project," which aims to "identify and explain persistent inequalities" (Haslanger 2000, 36). This is contrasted with the "ameliorative project," which aims at providing conceptual "tools in the fight against injustice" (ibid., 36). Despite their distinct goals, the two are not unrelated, as Haslanger (2006, 96) rightly admits that these "projects cannot, of course, be kept entirely distinct" and that "we might want to know the 'folk theory' of race (and other categories) in order to engage in hermeneutical deliberation about what we have meant and should mean" (2010, 180). Not only are the projects intertwined, but Haslanger affirms that continuity between concepts coming from the different projects is conceivably the final goal: "[W]e might hope that through reflection and discussion we could come to the point where (a) the concept we take ourselves to be employing, (b) the concept that best captures the type we are concerned with, and (c) the type we ought to be concerned with coincide. In such cases the conceptual, descriptive, and ameliorative projects yield the same concept" (Haslanger 2006, 96).

Accordingly, the success of Haslanger's definitions (with regards to the descriptive *and* ameliorative aims) requires that the folk have concepts of race and gender that extend beyond mere color and sex concepts. Saul's critique questions the plausibility of this (more on this below). Notably, Herman Cappelen endorses Saul's criticisms but takes it one step further by suggesting those difficulties are insurmountable: "What Haslanger's account needs, but doesn't provide, is a procedure for identifying 'the phenomena' that are being organized and explained. Haslanger has never responded to Saul's objection and my bold conjecture is this: *there simply isn't a good way to identify 'the phenomenon'* except disquotationally" (Cappelen 2018, 184; my emphasis). Saul is not as pessimistic as Cappelen, but surprisingly concludes there is no more work for the philosopher to do: "It may well be that Haslanger's terminology is so useful that this usefulness outweighs the ways in which I have worried that they might be counterproductive. But we need to know more in order to be able to evaluate this . . . So, in the end, it seems to me that what really needs to be done is some careful empirical work on the most effective vocabulary for combating racism and sexism. And this takes us outside the purview of philosophy" (Saul 2006, 142).

Saul and Cappelen are both mistaken. First, there is a good way to identify the phenomena that are the target of CE projects — by doing experimental philosophy! Recall that the positive program is aimed precisely at that goal; those philosophers are "motivated to explore intuitions experimentally because they think that by doing so they can do a better job of conceptual analysis" (Stich and Tobia, 2016). And as mentioned in the introduction, the use of experimental philosophy in methods of explication has already been advocated for by Shepard and Justus (2015), Pinder (2017), and Schupbach (2017). Second, the experimental methods needed to further the analysis of Haslanger's proposed definitions fall *inside* the purview of philosophy, and especially of CE. Outside the argument given in this paper, one can simply look at the activities of other philosophers. For example, Paul Griffiths's empirical work on genetic essentialist thinking, Josh Knobe's studies on

the factors influencing our moral judgments, Walter Sinnott-Armstrong's studies on moral cognition and decision-making, and Edouard Machery's investigations of knowledge attribution across cultures effectively rebut this dogmatic philosophy and science binary.

Crucially, as explained in section three, increasing precision usually increases measurability and testability. Perhaps if Haslanger had included precision as a guide at the start of her engineering project, the empirical difficulties highlighted by Saul could have been avoided or accounted for.[13] Fortunately for Haslanger, there is a sizable literature on race and gender concepts that provides the resources to rebut Saul's criticisms. One example, in particular, ought to convey the importance and relevance of that research. As mentioned above, Saul's critique casts serious doubt on whether Haslanger's ameliorative concepts represent the operative race and gender concepts of ordinary speakers. Again, this requires that the folk distinguish color from race concepts and sex from gender concepts, but Saul boldly conjectures the folk do not have anything resembling the kind of concepts Haslanger's ameliorative definitions presume:

> It seems to me that there is no good reason to even attribute gender or race concepts to ordinary speakers. Ordinary speakers do, of course, have concepts associated with the terms 'woman' and 'Black', but there is no reason to take these to be gender and race concepts — that is, to take them to be social correlates of sex and colour concepts. Moreover, it does not seem as though ordinary speakers have a tendency to organize the world (linguistically or otherwise) according to categories that really look like gender and race categories — rather than sex and colour categories, or some muddled amalgams. Why should we, then, attribute race and gender concepts to ordinary speakers? I'd like to tentatively suggest that we should not do so, although properly establishing this would require more argument than I am in a position to make here. If this is right, then it makes no sense to even ask whether Haslanger's analyses of race and gender concepts are revisionary with respect to ordinary ones: there are no ordinary race and gender concepts. (Saul 2006, 133)

Saul argues that ordinary speakers do not distinguish between color/race and sex/gender in the way that would track Haslanger's proposed race and gender concepts, but this claim is

[13] Despite Saul's criticisms, it's noteworthy that the exactness of Haslanger's definitions reflect an increase in precision that, as Brigandt and Rosario (2020, 104) note, defeated "prior scepticism about the possibility of putting forward a coherent concept of gender."

demonstrably false. Numerous studies on race and gender categorization, stereotype bias, self-concepts, and identity make obvious the fact that ordinary people have race and gender concepts that are more robust than mere associations with sex and color categories. In fact, several studies have actually investigated the content of racial concepts among the folk (see Martin and Parker 1995; Condit et al. 2003, 2004; Dubriwny et al. 2004; Jayaratne et al. 2006; Glasscow et al. 2009). Taken together, the results indicate that, in general, race is conceived of as a complex combination of ancestral, phenotypic, and social factors. In their research on focus groups, for example, Dubriwny et al. (2004, p. 187) explicitly asked participants, "What do you think is generally meant when people use the term 'race'?" That question was followed by others to discern whether participants were defining race by culture, geography, heredity or genetics, color, and religion. The results demonstrated that participants thought of race as multifactorial, that is, containing a mix of genetic, geographic, cultural, and socially constructed features. Similarly, Shulman and Glasgow (in press) administered a questionnaire that directly asked adults whether they think race is real or merely imagined, and then asked the realists to select whether people's races are determined by the way they look (classified as 'biological'), their social ties ('social'), their personality or abilities ('psychological'), or some combination thereof. Only about half of realists chose biology alone, and another 23% chose biology in combination with one or both of the other two determinants. Twenty-one percent chose a purely social conception of race. These data, along with the results of Dubriwny et al., challenge both the claim that race-thinking is univocally biological (or univocally social) and the conventional wisdom that race-thinking uniformly subscribes to essentialism and the one-drop rule[14] (Glasscow et al. 2009, 18).

[14] The 'one-drop rule' refers to the idea that race is typically thought of in terms of biological essences, such that if a person has even one Black ancestor, he or she is Black, for example.

The takeaway is that empirical work on race and gender is crucial to philosophical investigations on said topics. In a recent comparison between Carnapian Explication and Ameliorative Analysis, Dutilh Novaes (2018, 1027) notes that the precision criterion "has no immediate counterpart in Haslanger's framework" and suggests that "[t]his is, in turn, where practitioners of the ameliorative method may benefit from engagement with the techniques employed by Carnapian explicators." Exactly. My analysis of the debate between Saul and Haslanger demonstrates the benefit of including precision in ameliorative projects. In sum, given the epistemic rewards that typically come from increases in precision, conceptual engineers would benefit from making precision a methodological standard to guide future CE projects.[15]

2.5: Conclusion

I first argued that given the stated goals of conceptual engineers, empirical work is required in order to evaluate the success of CE projects. Second, I argued that precision should be a methodological standard for CE because of its utility in generating concepts amenable to the empirical methods required to evaluate successful CE projects. Finally, I demonstrated how ignoring precision can interfere with philosophical progress by highlighting where lack of precision left Haslanger open to Saul's critique. The upshot, however, was that this also revealed that Haslanger has the resources to rebut Saul's criticisms. To conclude, I will briefly mention how the methodological analysis presented here relates to a different methodology and what implications follow.

The so-called 'method of cases' has become prominent in discussions of philosophical methodology, especially in light of what has been dubbed the "experimenters challenge" (Nado 2015). The method of cases refers to the traditional practice of presenting cases (thought

[15] Haslanger has recently endorsed the interdisciplinary approach being argued for here. Haslanger (2020, 7) asserts that "the social critic embraces the normative dimension of philosophical theorizing, and also relies crucially on empirical research."

experiments) which describe actual—but more often hypothetical—situations and take our intuitions about the respective cases as evidence for what facts hold in the situation. Those intuition inferred facts are then brought to bear on philosophical controversies (Machery 2017). The experimenter's challenge refers to the large body of empirical evidence (some of which was summarized in this paper) that raises serious worries about the epistemic import of intuitions, thus undermining the method of cases. The argument presented here is not directly aimed at the method of cases debate, but it is obviously sympathetic to the experimenter's challenge. That said, the scope of my argument pertains to CE specifically. The argument should not be taken as purporting to establish an empirical methodology for philosophical inquiry *tout court*. This analysis leaves open the possibility that other methodologies may be fruitful for non-CE philosophical projects, a conclusion which is in line with Carnap's cooperative attitude: "We all agree that it is important that good analytic work on philosophical problems be performed. Everyone may do this according to the method which seems to be the most promising to him. The future will show which of the two methods, or which of the many varieties of each, or which combinations of both, furnishes the best results" (Carnap 1963, 939).

Lastly, it's worth mentioning that the argument has implications for the controversy surrounding what concepts are. Given that our concepts need to be empirically tractable in order to gauge their success, a naturalized account of concepts would be best suited for the purposes of CE. One option from Machery (2009) holds that concepts are psychological entities. Another option is the causal theory of conceptual content from Prinz (2002). A third is the theory-theory view, which claims that concepts are relational in the same way the content of scientific terms are specified by the role they play within a particular scientific theory. Analyzing these possibilities is outside the purview of this paper. I mention them only to highlight the interesting implications the methodological argument here has for ontology.

CHAPTER 3: SEX, SELECTION, AND SIMPLICITY

Abstract

This chapter does two things: first, it presents a case study that demonstrates the importance of not conflating explicanda and explicata. Second, it uses the experimental methods argued for in chapter one to support abandoning the concept of parsimony. Parsimony can be defined in a variety of ways against a variety of objectives. I examine a recent debate in behavioral ecology where it appears biologists are conflating two concepts of parsimony— quantitative and qualitative. I argue that failing to distinguish these two notions leaves the debate at a standstill and renders parsimony's supposed argumentative force impotent. Then I apply Sober's likelihood formulation to the respective hypotheses to determine if there is any epistemic role for parsimony in this debate. I conclude by arguing for an eliminativist stance towards parsimony by appealing to original X-phi research I conducted at Florida State University.

3.1 Introduction

All else being equal, we ought to prefer a simple theory or hypothesis over a more complex one—or so the principle of parsimony proclaims.[16] Without some precisification, this general statement is too elusive to evaluate its epistemic significance. The problem is what, exactly, does simple mean? Does it merely refer to less stuff? That is, minimizing the number of processes, entities, causes, or assumptions postulated. Or does simplicity refer to fewer kinds of stuff? In other

[16] Parsimony and simplicity are used interchangeably throughout this chapter.

words, minimizing the types of processes, entities, causes, or assumptions postulated. The former is referred to as quantitative parsimony, and the latter qualitative parsimony.

A non-scientific example helps elucidate the difference: Imagine you are comparing two dresser drawers in your room. In the first drawer, you have five items: one pair of shorts, two socks, and two shirts. In the second drawer, you have ten items and they are all socks. The first drawer is quantitatively parsimonious because the number of clothing items is fewer than —5 vs. 10. The second drawer is qualitatively parsimonious. The first drawer has three types of garments, but there is only one type in the second drawer: socks!

Assuming parsimony can be defined more precisely (perhaps via the definitions above), the problem of justification remains. As Sober candidly remarks, "It is obvious that simple theories may be beautiful and easy to remember and understand. The hard problem is to explain why the fact that one theory is simpler than another tells you anything about the way the world is" (Sober 2015, 2).

If there is going to be a solution to the 'hard problem' it is not going to be an a priori one. Following Sober and others (Popper 1959; Jeffreys 1961; W. v. Quine 1963; Lange 1995; van Fraassen 1980; Goodman 1955; Baker 2007), I maintain that parsimony is not a final scientific end in itself and its status as a theoretical virtue is conditional. If parsimony is relevant and virtuous, it's only contingently so: relative to a particular research context and against a set of specific a posteriori background assumptions. Much of the recent work on theoretical virtues has focused on determining what those subject-matter specific conditions are. This chapter is not contributing to that project: I am not attempting to identify any circumstances where parsimony inferences are justified. Instead, the primary goal is to seed doubt about the relevance of parsimony in the first place.

Section 2 presents a case study from behavioral ecology where opposing sides of a debate each appeal to parsimony as a reason to favor their hypothesis over the rival. But, at least in some

cases, the competing hypotheses are more parsimonious in disparate ways. Some biologists use the quantitative concept of parsimony, while others use the qualitative: sometimes they use both! Such trite appeals to parsimony call into question its supposed cogency. Of course, regardless of its confounding usage, it remains open whether epistemic significance is being correctly picked out by parsimony. Section 3 applies Sober's (2015) likelihood framework to the debate considered in section 2 and argues that while a Bayesian approach to parsimony is available, it does not adjudicate between the two hypotheses in the debate. Section 4 presents original experimental findings that strongly indicate that the term "simple" produces a framing effect and argues that this data supports an eliminativist stance towards parsimony in scientific reasoning.

3.2 Sex

It has been widely observed that there are behavioral changes associated with the menstrual cycle. In women, preference for male-specific traits that may be markers of mate quality (like masculine face shapes and bodies and deep voices) increases. Further, women's perceived physical attractiveness appears to vary across the menstrual cycle. For example, between-individual attractiveness in faces is plausibly mediated by factors like symmetry, sexual dimorphism, averageness, and overall health, which vary across the menstrual cycle. Additionally, although the relative significance of these traits likely varies across cultures, most researchers agree that males' preference for facial femininity, gynoid fat distribution, waist-to-hip ratio, vocal characteristics, and breast shape are preferences that might reveal variation in mate quality (see Gildersleeve, G. Haselton, and Fales 2014 for an overview). Importantly, shifts in these preferences also occur within individuals. For example, facial photographs of women during the follicular phase receive higher attractiveness ratings than photographs of the same woman during the luteal phase.[17] These

[17] The follicular phase refers to the time beginning on the first day of menstruation and ending with ovulation during which the follicles in the ovary mature. The luteal phase starts after ovulation and ends with the beginning of menstruation.

correlations associated with the menstrual cycle are also observed for voices, axillary, and vaginal odors (Gildersleeve, G. Haselton, and Fales 2014).

Given the pervasiveness of these trends, it has been assumed that the mechanisms for discriminating these traits are adaptations (Thornhill and Gangestad 2008). But because the changes associated with the traits are so subtle, most researchers conclude that cues for ovulation are adaptively concealed. Thus, two adaptations are being hypothesized here: first, an adaptation in females for the concealment of ovulation, and second, a distinct psychological adaptation in males to detect cues of ovulation. For females, the supposed benefit of concealed ovulation is the opportunity to pursue a mixed-mating strategy, that is, extra-pair copulation with high-quality mates when conception is possible. For males, the ability to detect current fertility may be adaptive because it could increase the chances of conceptive sex and/or be a signal to intensify mate-guarding behaviors around ovulation to avoid the costs of cuckoldry (Thornhill and Gangestad 2008; M. Haselton and Gildersleeve 2011). The first proposed adaptation, that female mate choice is influenced by increased attraction to masculine traits, is called the "human estrus" hypothesis. The second adaption, that men can detect ovulation by perceiving females' increase in sexual attractiveness during the periovulatory phase of their menstrual cycle, is called the "ovulatory shift" hypothesis.

Despite their intuitive plausibility, Havliček et al. 2015b (hereafter HCBKR) argue in a recent paper that the evidence for these putative adaptations is more parsimoniously explained as by-products. They argue that the discernment of ovulation by males and the cyclical changes in females' mate preferences are both by-products of hormone-dependent between-individual variation in cues and preferences. In other words, instead of one adaptation for within-cycle mate preference variation and a second adaption for discriminating ovarian hormone-dependent individual quality, selection acting on between individual expression produces these traits as an epiphenomenon.

HCBKR gives four reasons to reject the two-adaptation account. First, a flaw with the ovulatory shift hypothesis is that ovulatory changes in females' physical appearance are almost impossible to measure with our normal capacities. Even more sophisticated techniques, like detailed acoustic analyses of females' voices, cannot unequivocally identify the different ovulatory phases (HCBKR, 1251). This suggests that men are not able to reliably determine when a woman is likely to ovulate, and thus any shifts in attractiveness (if such shifts exist) are biologically insignificant.

Second, for the human estrus hypothesis, the increased preference for masculine traits during the follicular phase of the menstrual cycle is very small and probably does not have any relevance to human mate choice behavior in the real world. In particular, HCBKR points out that the studies on menstrual cycle-related attractiveness almost always involve participants that are unknown to each other. They also note that "studies in other animals suggest that the limits of detection of variability in traits such as symmetry, assessed in laboratory-measured tests, renders such traits ineffective for discriminating natural variation" (ibid., 1252).

Third, another issue with the human estrus hypothesis is that the two latest meta-analyses report inconsistent results regarding the robustness of the phenomenon. Proponents of the two-adaptation explanation have tried to account for the discrepancy by pointing to differences in the methods used in each analysis. But on the by-product account the discrepancy is expected:

> The most interesting aspect of the current meta-analyses, in our [HCBKR] view, is the size of the effects found and the variability in these across different traits; it seems that these do not just represent variability in the rigor of the methods and the definitions employed in such studies, but tell us something about the phenomenon itself . . . The variable and elusive nature of the phenomenon is precisely what one would expect from a by-product explanation: that is, variability is not the consequence of experimental/measurement error and noise, but an inherent characteristic of the phenomenon under study. (ibid., 1255)

Fourth, the two-adaptation explanation assumes that at some point in modern human's evolutionary lineage selection favored females with comparatively less obvious cyclic changes because of the potential advantages for paternity confusion. But HCBKR claims that it is uncertain

whether there was selection for concealment in the human lineage based on current phylogenetic studies and asserts the by-product hypothesis "does not require such an evolutionary step" (ibid., 1257). Given these considerations, HCBKR conclude that:

> Although one could postulate two independent adaptations — one for within cycle variation in preferences and a second being sensitive to ovarian hormone-dependent individual quality — a more parsimonious explanation is that selection acting on between-individual expression also produces cyclical variation as an epiphenomenon. [ ] According to our view, within-individual changes in women's attractiveness and preferences across or between cycles are most likely by-products, or perceptual spandrels, rather than specific and independent adaptations. (ibid., 1256; my emphasis)

Notice the use of parsimony here and recall the aforementioned distinction between qualitative and quantitative parsimony. We can stipulate more specifically that HCBKR take their explanation to be *quantitatively* parsimonious; the phenomena in question are by-products of one common adaptation, whereas the human estrus and ovulatory shift hypothesis posit two independent adaptations.

This proposed by-product explanation has been met with mixed reviews. A glance at the titles of responses indicates how central parsimony appears to be to this disagreement. For example, "Is the perspective of Havliček et al. really new (or truly parsimonious)?" and "Greater precision, not parsimony, is the key to testing the periovulation spandrel hypothesis." Next, we will closely examine some of these responses which reveal the conflated and confused use of parsimony.

3.2.1 Responses to HCBKR

Some researchers are quite convinced by the by-product proposal: "surely it is time to acknowledge that the 'ovulatory shift' and 'human estrus' hypotheses are flawed, and to accept that they should be discarded" (Dixson 2015, 1261). But this positive assessment is not unequivocally shared. For example, Gangestad and Grebe claim that HCBKR "vastly overstates the case that researchers have neglected particular kinds of explanations (e.g., by-product hypotheses) and

insufficiently acknowledge substantial empirical and theoretical challenges their proposals face" (Gangestad and Grebe 2015, 1262).

First, consider the response, "A new (or truly parsimonious) perspective? A comment on Havlíčet al." (Haselton 2015). Haselton presents two arguments to undercut the by-product explanation's simplicity. To start, recall that HCBKR's third criticism was about the small and variable effect sizes reported in two meta-analyses regarding cycle-shifts in female mate preferences. Against this conjecture, Haselton argues that the small and variable effect sizes are likely due to an imprecise method for assessing ovulation. Usually, fertility is determined via participants' recalled menstrual onset. However, such methods "have modest validity and attenuate estimates of effect sizes by 50% or more. . . Corrections for low validity indicate that true effect sizes may be in the medium to large range" (Haselton 2015, 1264). This suggests the supposed inconsistency between the two meta-analyses is not correct. According to Haselton, a re-analysis of the Wood et al. (2014) data produced evidence of cycle shifts consistent with those initially documented by Gildersleeve et al. (2014). Further, the effects measured in both meta-analyses were evaluated for short-term mating preferences and not when females evaluate males as long-term mates. Thus, the by-product explanation can't maintain its claim to parsimony because it saddles us with a new question: If shifts in women's mate preferences across the menstrual cycle are by-products of more general hormone-dependent effects that facilitate mate choice, as HCBKR claim, then why are these effects absent when evaluating long-term mates?

Haselton's second argument pertains to the ovulatory shift hypothesis. Haselton argues there are studies that suggest males can detect subtle ovulation ques, but "only in the context of romantic partnerships characterized by frequent contact." (ibid., 1264) She agrees with HCBKR that the effects of males' preference for attractive features between females that index general fertility are likely to be higher than males' detection of ovulation cues in long term partnerships. However, she

contends that just because the former set of effects is likely to be larger than the latter, "it does not logically follow that the latter do not exist." (ibid., 1264) This logical point, coupled with the criticism regarding long-term vs. short-term mating strategies, leads Haselton to conclude that "It is premature, then, to claim that the between [individual] perspective is the most 'parsimonious' explanation for estrous-like shifts." (ibid., 1264)

Neither of these arguments obviously subverts the by-product explanation's claim to parsimony. Haselton's first argument — that a new explanation for the absence of a correlation between preference shifts and menstrual cycles in evaluating long-term partners is needed — only works if that explanation required positing more adaptations. The problem with this argument is that it is entirely unclear the answer to the question "why is there not a correlation between mating preferences and the menstrual cycle in evaluating long-term partners?" would require (or likely require) positing additional adaptations. Adaptations are a historical concept, not to be confused with current utility. A selective history might account for a trait's existence in the population without being currently adaptive. (Sober 2000, 85-86) Without evidence that the correlational discrepancy we observe *now* was present in the evolutionary *past* when the trait in question supposedly evolved, this route to undermining the by-product's simplicity does not seem promising. Another possibility is that the answer to Haselton's question would require positing other non-selective forces. Accordingly, HCBKR's explanation would contain one adaptation + another mechanism of evolution. Notice, however, this would not affect the relevant notion of quantitative simplicity, which HCBKR specifically applies to the number of adaptations. Having one adaptation + something else would change its qualitative parsimony, the number of kinds of processes posited. Any suggestion that the by-product account is not *truly* parsimonious in light of that possible

consideration incorrectly assumes the qualitative notion of parsimony is the only one salient to scientific reasoning.[18]

Haselton's second argument was that an adaptation for ovulation detection in men is not logically excluded from the evidence cited by HCBKR in favor of the by-product explanation. There is no disagreement on this conceptual point. Why we should think this logical possibility tells us anything about the reality of the biological world is not obvious. At most, the logical point indicates that the current body of data underdetermines the competing hypotheses. As presented, Haselton's arguments do not subvert the quantitative parsimoniousness of the by-product explanation. Let's examine another response.

Gangestad and Grebe (2015) argue the by-product explanation is not phylogenetically parsimonious. They begin by asserting data from Thornhill and Gangestad (2008) shows that many male non-human primates have adaptations for ovulation detection using scent cues and that adaptive estrous cycle shifts of sexual preferences in female non-human primates were found by Dixson (2012). Gangestad and Grebe then ask:

> How does a species that has no hormone-dependent adaptations based on within-cycle variation (as HCBKR propose for humans) evolve from an ancestral species that does, yet retain these hormone-dependent shifts as "by-product"? . . . [HCBKR] offer no vision of an evolutionary process whereby within-cycle shifts changed from adaptations for within cycle variation to byproducts of between-woman variation. For example, were within-cycle variations lost at some point in humans, only to reappear as by-products? Such a phylogenetic scenario strikes us as even more "elaborate," but less likely, than the scenarios HCBKR reject. (Gangestad and Grebe 2015, 1263; my emphasis)

In response, Havliček et al. 2015a claim Gangestad and Grebe's argument only gains traction under the assumption that female primate sexuality is tightly linked to the menstrual cycle and thus to ovarian hormones. However, HCBKR argues such an assumption should be rejected. That

[18] See Nolan 1997, "Quantitative Parsimony."

argument is not relevant for the purposes of this analysis. The interesting element is what HCBKR claims follow from rejecting Gangestad and Grebe's assumption: "if this premise is incorrect, then the question of parsimony becomes less relevant (note also that evolutionary processes are not always parsimonious)." (HCBKR 2015b, 1266)

This is where things start to get muddled. Phylogenetic parsimony *typically* refers to the number of steps required to generate the observed variation from common ancestral sequences. Adaptationism, specifically, has to do with the "power" of natural selection. That is, whether the causal role of selection is so strong most of the observed variation we see can be explained by models where natural selection is described, and non-selective processes are omitted. (See chapter five in Sober 2000) In other words, adaptationism is a thesis about the mechanism involved in each evolutionary step, and nothing about that thesis by itself entails anything about the number of steps that took place. It is not obvious which notion Gangestad and Grebe had in mind when claiming the phylogenetic scenario on HCBKR's view would be "more elaborate" than the two-adaptation account. Based on their question, "were within cycle variations lost at some point in humans, only to reappear as by-products?" it seems like they understand parsimony as minimizing the number of character changes that took place. But this is not the notion of parsimony HCBKR originally attributed to their by-product account, which minimized the number of observed similarities that cannot be explained by natural selection. Thus, it appears there are at least two different ways in which an evolutionary explanation can be quantitatively parsimonious.

My aim is not to take a stand on which hypothesis is correct. What I hope to have shown is how lack of clarity has severely obscured this debate. No one in the literature we've considered explicitly states how they are defining parsimony. And it is unfortunate that various notions of simplicity are picked out by the same vague word 'parsimony.' Without being more precise, it is

difficult, if not impossible, to evaluate a particular parsimony assignment. Notably, this conclusion is strikingly similar to Carnap's diagnosis of the problem of probability:

> Von Mises and Jeffrey both assert that there is only one concept of probability . . . This is primarily due to the unfortunate fact that both concepts are designated by the same familiar, but ambiguous word 'probability' . . . The first maintains that this concept is probability 2 . . . the second puts it the other way around; and neither has anything but ironical remarks for the concept proposed by the other . . .These mutual rejections are often formulated in rather strong terms. This whole controversy seems to me futile and unnecessary. The two sides start from different explicanda, and both are right in maintaining the scientific importance of the concepts chosen by them as explicanda. (Carnap 1945, 518)

Despite the similarity of these analyses, unlike the case of probability, the scientific importance of parsimony is far from established. We turn to this issue next.

3.3 Selection

In the previous section, two distinct definitions of parsimony were specified and used to analyze the peri-ovulatory debate. But I didn't comment on whether the quantitative and qualitative notions are successful explications of the parsimony concept to begin with. Unfortunately, there is no empirical evidence that connects these vague notions of simplicity to a theory's plausibility, reliability, or credibility. (Quine 1963) This empirical discrepancy catalyzed the search for probabilistic justifications for parsimony by looking at various statistical inference frameworks such as frequentist, Bayesian, likelihood, and multi-model frameworks. (Sober 2015) I will focus exclusively on Bayesian approaches.

If our explicative goal is a concept of parsimony that is consistent with a Bayesian epistemology, then simpler theories need to have higher priors than more complex ones, simpler theories need to have a higher likelihood or both. Traditional Bayesians favor prior probabilities and try to justify parsimony by assigning simpler theories higher priors (see Jeffrey 1965). Likelihoodists decline to pursue this route because they claim there is no systematic connection between simplicity and higher priors (Sober 2015). I will take up the traditional Bayesian approach in the next section.

First, I argue a likelihood formulation of parsimony vindicates one of the explanations in the peri-ovulatory debate.

The connection between simplicity and likelihood is Reichenbach's principle of common causes. The basic idea is that a correlation between two events indicates that one event causes the other or that both events have a common cause. On the likelihood approach, evidence E favors a simpler hypothesis S over a more complex alternative C iff $P(E|S) > P(E|C)$. In other words, justifying parsimony in a specific context requires producing an argument that this inequality holds. Skipping over the mathematical details, Sober stipulates seven assumptions that must be met on the likelihood approach:

1. Common causes screen-off their effects.
2. The probabilities in the model are nonzero.
3. There is a positive correlation of the cause with each effect.
4. The separate causes are probabilistically independent of each other.
5. Separate causes together screen-off each effect from the other.
6. Each component cause screens-off the other.
7. The values of the parameters in the common cause model are equal to the values of the parameters used in the separate cause model.

Assumptions 1-3 pertain to common causes; they tell us that if there is a common cause between X and Y, then X and Y will be correlated. Assumptions 4-6 apply to separate causes; they tell us that X and Y should be uncorrelated. When we can assume that the processes generating the data are probabilistically independent and identically distributed, then assumptions 1-6 can be enough to conclude (pending the observational data) that:

> Pr(X and Y are positively associated | the CC model defined by Assumptions 1-3) > Pr(X and Y are positively associated | the SC model defined by Assumptions 4-6)

When the processes are not independent and identically distributed then Assumption 7 connects the two models. (Sober 2015, 111-114) When a testable problem meets these assumptions, then explanations that posit a common cause will have a higher likelihood than a competing separate cause explanation. Thus, common-cause explanations can be preferred on Bayesian grounds because the law of likelihood entails that the common-cause explanations will have a greater posterior probability.

This is my proposal: Because the by-product explanation claims the respective explanandum are epiphenomena of one common adaptation, we can think of it as a common-cause explanation (CC). The alternative explanation posits two independent adaptations in addition to the adaptation that undergirds the CC. Accordingly, the two independent adaptation explanation is a separate cause explanation (SC). Recall, the phenomena to be explained are (1) preference shifts across or between cycles for male-specific traits that might be markers of mate quality and (2) within individual changes in women's attractiveness that are supposedly detectable by males. Call the former trait X and the latter trait Y. Assumption 7 from above can be satisfied by specifying different character states (Ti) for X and Y by the features that characterize attractiveness in males and females, like face and body shape, symmetry, sexual dimorphism, voices, and axillary odors, etc. Thus, our data set will include several different observations, one for each of the characters specified (Ti...Tn). The crucial question then is whether X and Y are correlated. The by-product account predicts changes to both X and Y when general between-female hormone differences are manipulated. This is what most studies since HCBKR (2015) have tried to investigate, and the results are a mess—with half as many studies reporting a correlation as there are reporting none or the opposite correlation expected.

So, given the current body of evidence, the simplicity of HCBKR's by-product account does not give us reason to prefer it over the two-adaptation explanation. That could change as new evidence rolls in. All I have intended to show is that a likelihood formulation of the problem is

possible, thereby establishing at least one epistemically relevant sense of parsimony in the peri-ovulatory debate.

3.4 Simplicity

Three different explications of the term parsimony have been put on the table: a quantitative, a qualitative, and a likelihood formulation. The quantitative and qualitative versions helped elucidate the use of the term in the peri-ovulatory debate, but neither definition had been made precise enough to clearly demonstrate their epistemic significance. On the other hand, the likelihood formulation has clear epistemic import. Notice, however, the likelihood version was absent from the biology literature examined. I will conclude by arguing for a general eliminativist stance towards parsimony in scientific reasoning.

The case study presented here is about evolutionary explanations. In particular, what explains the existence of certain behavioral traits in the present population. The only route to take in answering ontological questions like this is by looking at the data: what does the evidence say? (Sober 2010) My hypothesis is that the word 'parsimony' draws attention away from the purely evidential considerations needed to support an evolutionary explanation—a worry also shared by some biologists recently.[19] There are several reasons why I think this hypothesis is correct. First, analyzing how parsimony was used in the peri-ovulatory debate revealed that the term is unnecessarily confounding. Beneath the violent back and forth over which explanation is 'truly' parsimonious are disputes about methods of data collection, the structure of data analysis, and assumptions about the nature of evolution, for example. These are legitimate points of contention that can reasonably affect our evaluation of the candidate hypotheses, regardless of how 'parsimonious' the explanations may be.

[19] For example, Coelho, Diniz-Filho, and Rangel 2019 claims "The principle of parsimony can be dangerously used as a silver bullet to solve any argument when contrasting theories, hypothesis and models." (974)

Second, psychologists have recently started studying the role of simplicity in evaluating competing causal explanations and the findings lend some support to this hypothesis. (Lombrozo 2007; Pacer and Lombrozo 2017; Blanchard, Lombrozo, and Nichols 2018) The experimental results indicate that, in general, simpler explanations are not only preferred but also judged more likely than complex ones *given the same evidence*. Specifically, simplicity considerations affected the assessment of an explanation's prior probability by a factor of four. (Lombrozo 2007) Thus, for a simple and a complex hypothesis with the same initial evidential support, a disproportionate amount of additional evidence is required for someone to forgo the simpler explanation for the more complex one. Moreover, the prevalence of a cause invoked in a simple explanation is consistently overestimated when an unlikely— but simple —explanation is endorsed.

Finally, I collaborated with Jack Justus and Nick Byrd at Florida State University to replicate and extend the results from Lombrozo 2007. Lombrozo uses two metrics to measure simplicity. The first is node simplicity: the number of causes an explanation invokes; the second is root simplicity: the number of unexplained causes an explanation invokes. Lombrozo's studies ask participants to evaluate a simple explanation against more complex alternatives. Our analysis targeted something different: whether the word 'simple' and underlying concept produces a framing effect. In other words, whether merely characterizing scientific explanations as simple affects our explanatory judgments regardless of whether an explanation is simple according to some specified metric.

In order to gain insight into how simplicity and probabilistic considerations interact, Lombrozo (2007) presented participants with the following vignette:

> There is a population of 750 aliens that live on planet Zorg. You are a doctor trying to understand an alien's medical problem. The alien, Trackie, has two symptoms: Trackie's *minttels are sore* and Trackie has developed *purple spots*.
>
> - *Tritchet's syndrome* always causes both *sore minttels* and *purple spots*. You know that Tritchet's syndrome is present in about 50 of the aliens on Zorg.
> - *Morad's disease* always causes *sore minttels*, but the disease never

causes *purple spots*. You know that Morad's disease is present in about 197 of the aliens on Zorg.

- When an alien has a *Humel infection*, that alien will always develop *purple spots*, but the infection will never cause *sore minttels*. You know that Humel infection is present in about 190 of the aliens on Zorg.
- Nothing else is known to cause an alien's *minttels to be sore* or the development of *purple spots*.

What do you think is the most satisfying explanation for the symptoms that Trackie is exhibiting?

(A) Trackie the alien has Tritchet's syndrome.
(B) Trackie the alien has Morad's disease.
(C) Trackie the alien has a Humel infection.
(D) Trackie the alien has Tritchet's syndrome and Morad's disease.
(E) Trackie the alien has Tritchet's syndrome and a Humel infection.
(F) Trackie the alien has Morad's disease and a Humel infection.

The symptoms could be explained by an appeal to a single disease (Tritchet's syndrome) or two diseases (Morad's disease and a Humel infection). Lombrozo modified the base rate frequencies for each disease across eight different conditions to represent a range of probabilistic support for the more complex (i.e., two-disease) explanation.

See figure 3.1.

$P(D_1):P(D_2 \text{ and } D_3)$	D_1	D_2	D_3
15:1	50	50	50
1:1	50	197	190
9:10	50	195	214
4:5	50	225	210
2:3	50	250	220
1:2	50	268	280
1:3	50	330	340
1:10	50	610	620

For each probability ratio, the corresponding number of aliens with each disease (from a population of 750) is indicated. Each probability ratio corresponds to a single baserate condition.

Figure 3.1: Reprinted from Lombrozo (2007,240)

The results of Lombrozo's experiment are represented in figure 2 below.

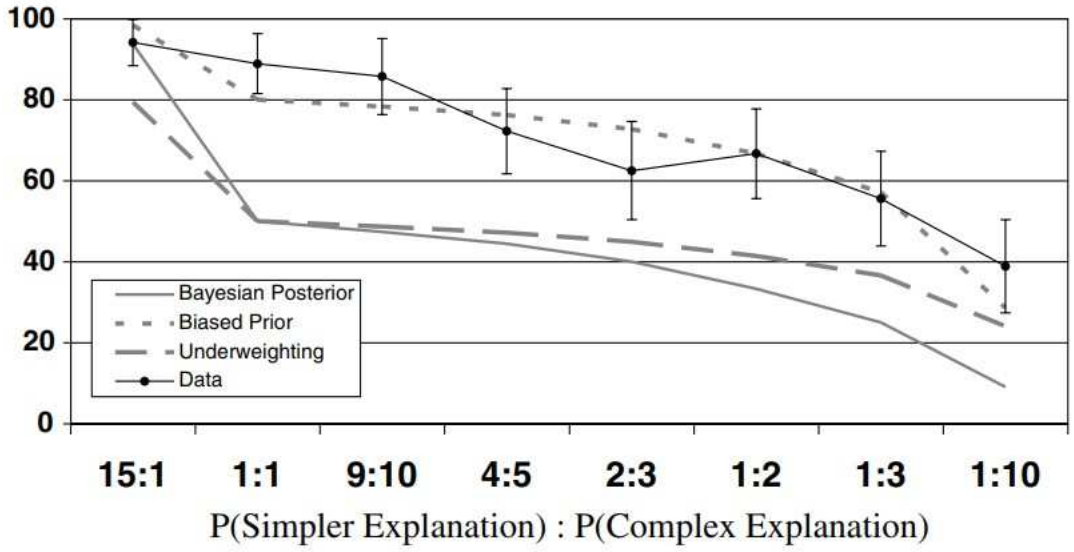

Figure 3.2: Reprinted from Lombrozo (2007, 241): the x-axis corresponds to the probability ratio of the simpler explanation to the more complex alternative. The y-axis corresponds to the percentage of participants selecting the simpler explanation. The data are indicated with black circles; the solid gray line presents the values that would be expected if the percentage of participants selecting the simpler explanation corresponded to the Bayesian posterior probability for the simpler explanation at the corresponding probability ratio ("Bayesian posterior"). Also illustrated are predictions corresponding to two ways in which the Bayesian calculation might be biased: by involving a prior probability favoring simpler explanations ("biased prior") or by under-weighting the relevance of the provided probability information ("underweighting").

As the figure shows, even when Morads and Humel were ten times more likely than Tritchets, over a third of participants selected Tritchets as the most satisfying. In our experiments, we replicated and performed our manipulation on the 1:1 and 1:3 base rate conditions from Lombrozo's original study. We recruited 438 participants who had at least a bachelor's degree from Mturk to complete the survey in exchange for payment. For each base rate there were three conditions.

Condition one replicated Lombrozo's vignette with no modifications. In order to test the hypothesis that merely using the word 'simple' affects explanatory judgments, we added a simplicity frame to the other conditions. Condition two added a simple frame to the simple explanation by including the sentence "The simplest explanation is that Trackie has Tritchet's syndrome." before the survey question: What do you think is the most satisfying explanation for the symptoms that Trackie is exhibiting? Condition three added a simple frame to the complex explanation by including

the sentence "The simplest explanation is that Trackie has Morad's disease and a Humel Infection." before the survey question. Our hypothesis predicts that the presence of the simple frame would cause participants to deviate from Lombrozo's results and in the direction of the explanation the frame was attached to (one disease or two diseases). Our findings suggest that the word "simple" does produce the framing effect our hypothesis predicts. The results are represented in figure 3 (below).

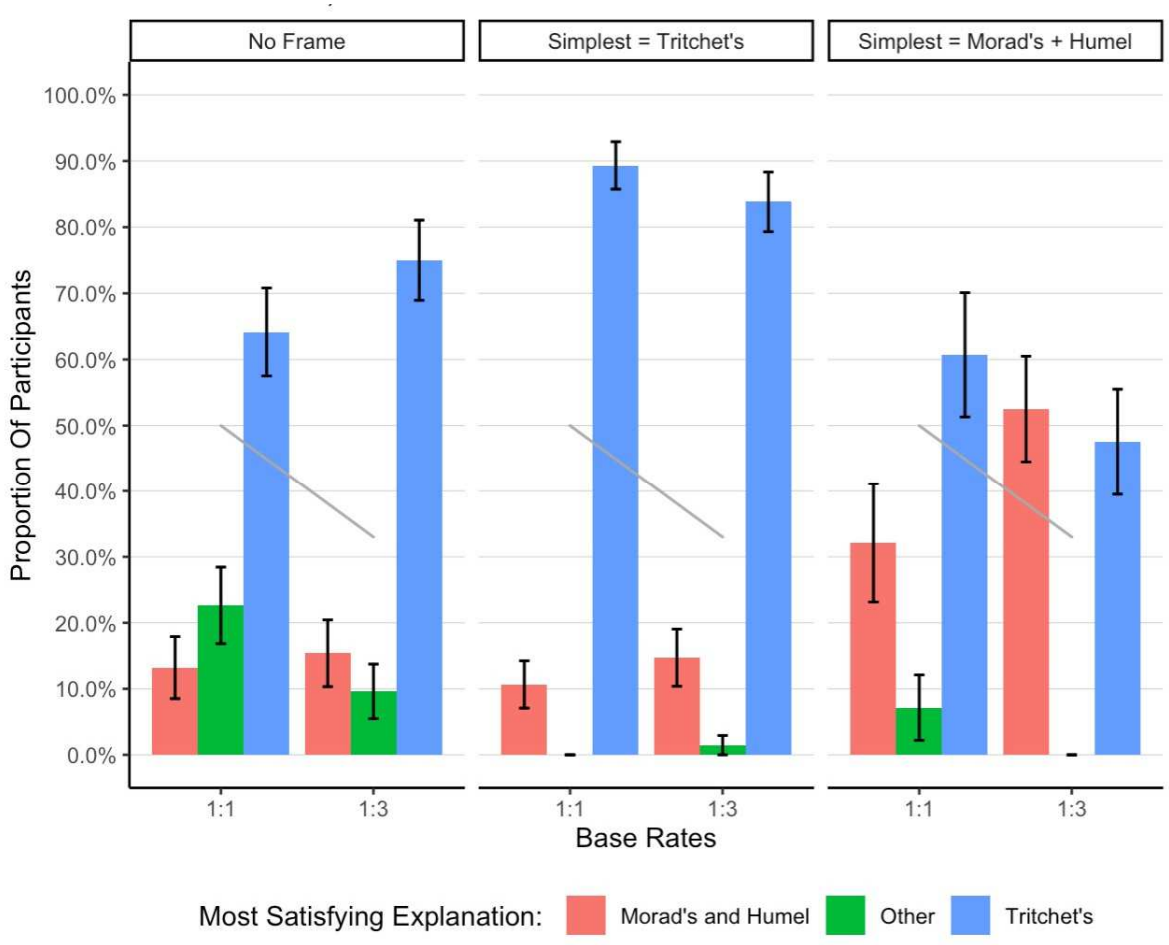

Figure 3.3: Explanatory preference for each frame by the base rate for participants that passed the comprehension check (N=438). Error bars represent a standard error from the mean. Solid grey lines represent Bayesian posteriors for the "simpler explanation" (a la Lombrozo 2007).

Non-overlapping error bars represent a significant effect. To test the effect of the simple manipulation on explanatory preference, ANOVA was used to compare the explanation choice in the frame vs. no frame condition for each base rate. Among those who passed the attention check,

the simple frame increased preference for Tricket's had a significant effect in the 1:1 base rate condition (p= 0.001294) but not in the 1:3 base rate condition (p= 0.2453). The effect on increased preference for Morad's and Humel's was marginally significant in the 1:1 base rate condition (p= 0.06883) and very significant in the 1:3 base rate condition (p= 0.0002054). We also compared the preference for Tricket's verses Morad's and Humel's for each base rate in the frame conditions. For both base rates there was a very significant effect, 1:1 (p= 0.007364) and 1:3 (p= 0.0001919). Finally, we analyzed whether perceived simplicity predicts perceived correctness across all explanation choices and the effect was very significant ($p < 0.0001$). See figure 4.

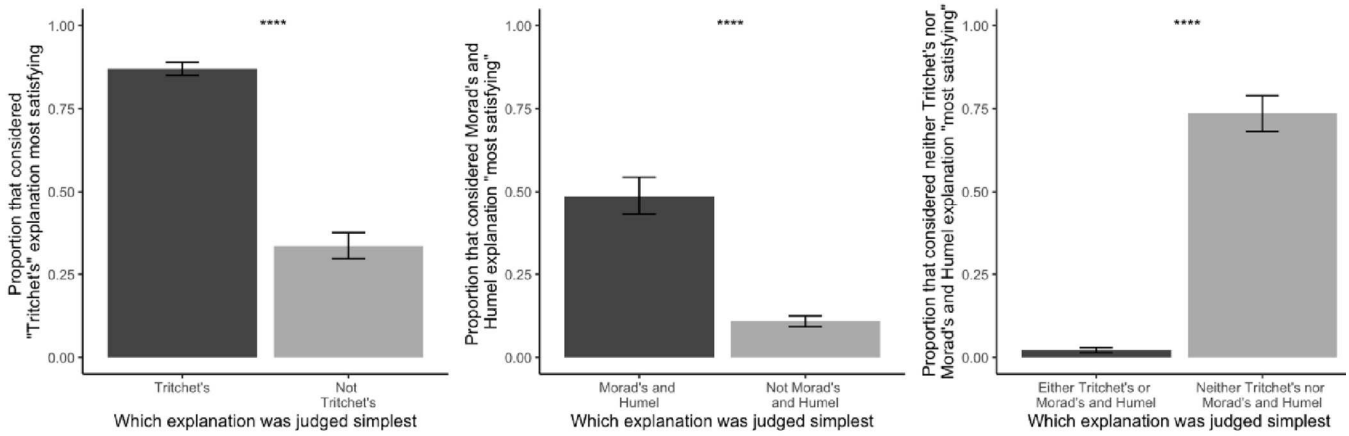

Figure 3.4: Predicting responses to "which explanation is most satisfying?" from responses to "which explanation is simplest?". (N=438)

3.5 Conclusion

Sober's (1990) analysis began by stating "respect for the results of science . . . leads one to assume that general principles of simplicity and parsimony must be justified" (E. Sober 1990, 90). Through a compelling example regarding the debate about group selection, Sober clearly demonstrates how claims of parsimony are actually covert appeals to local background assumptions. This led to his conclusion that there might be good reasons to 'razor Ockham's razor.' Later, his development of the likelihood formulation motivated him to retract this previous eliminativist stance. In light of the experimental results presented here, respect for the results of science seems to

require a return to eliminativism. Given the explicit epistemic aim parsimony supposedly serves in scientific practice, such a robust framing effect strongly indicates efforts to explicate the parsimony concept are all for naught. These attempts seem doomed to fruitlessness since merely invoking the term causes individuals to differentially (and non-evidentiarily) evaluate how well evidence supports hypotheses. This provides more than a compelling reason for razoring the razor in the context of evolutionary biology. It provides a data-driven and empirically justified rationale for eradicating the concept of parsimony from scientific theorizing in general.

CHAPTER 4: THE FUNCTION OF FRUITFULNESS

Abstract

In this chapter, I argue that Pinder's (2020) Relevant-Goals Account of fruitfulness is inadequate for the purposes of conceptual engineering. First, Pinder claims conceptual engineers need a concept of fruitfulness that extends beyond Carnap's conception. However, Pinder's case study fails to demonstrate that extension. Second, Pinder supplements the account with a view he calls Contextualism, which is supposed to specify which theoretical goals are relevant and provide substantive details for specific explications not derivable from the general account alone. Contextualism is on the right track, but I argue that Contextualism does not capture the pluralistic nature of conceptual engineering, whose goals are not always theoretical. To that end, I defend a pluralistic account of fruitfulness which is better suited for the diversity of conceptual engineer's goals and projects. Finally, I analyze a case study about explicating gender concepts which demonstrates the usefulness of the pluralist account.

4.1 Introduction

Conceptual engineering (CE) is a research program aimed at analyzing, refining, and potentially replacing concepts. The motivation behind CE comes from recognizing that oftentimes a concept fails to fulfill certain functions that researchers want the concept to do. Examples of this abound in science. For instance, a biologist might want a concept of the 'niche' that enables us to understand controls on a species' geographical range limits:

> In this article, after summarizing Hutchinson's niche concept and sketching empirical approaches to quantifying niches, I will point out limitations in standard articulations of the Hutchinsonian niche. Some arise from its focus on a species solely when it is rare, which can belie the impact of positive density dependence and feedback processes on population

> persistence; others arise from its relative neglect of temporal variation and spatial dynamics; and yet others arise from genetic variation and evolution in the niche itself. These reflections will lead to alternative and complementary niche definitions, enriching and extending Hutchinson's concept. (Holt 2009, 19659)

But CE isn't only concerned with scientific concepts. Social and political theorists and activists also engage in conceptual engineering. For example, Sally Haslanger states that her research project "seeks to identify what legitimate purposes we might have (if any) in categorizing people on the basis of race or gender, and to develop concepts that would help us achieve these ends." (Haslanger 2012, 366) More specifically, the goal "is to develop accounts of gender and race that will be effective tools in the fight against injustice" (Haslanger 2000, 36). Rawls also had an explicit practical agenda "to articulate a public conception of justice that all can live with who regard their person and their relation to society in a certain way . . . doing this may involve settling theoretical difficulties, [but] the practical social task is primary" (Rawls 1980, 519). Appeal to conceptual revisions are even found in more traditional philosophical debates within metaphysics and epistemology (Sharp 2013; Eklund 2020).

Of course, if we see the value in CE and are already engaged in those kinds of projects, then we ought to think about how exactly we should be doing it. What is a good methodology for CE? Most philosophers have taken Carnap's method of Explication as a jumping-off point. As Carnap describes it: "The task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second. We call the given concept (or the term used for it) the explicandum, and the exact concept proposed to take the place of the first (or the term proposed for it) the explicatum" (Carnap 1950, 2).

There are three basic steps to an explication. First, you must clarify—as much as possible—what the explicandum (the original concept) is. The meaning or function of the concept must be stated clearly to avoid confusion and ensure you are including the intended features in the explicated concept. The second step is to develop the explicatum (the new concept), which is usually given by

explicit definition. Third and finally, you should assess the success of the explicatum according to the following four criteria (1950, 7):

> (1) The explicatum is to be *similar* to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
>
> (2) The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an *exact* form, so as to introduce the explicatum into a well-connected system of scientific concepts.
>
> (3) The explicatum is to be a *fruitful* concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).
>
> (4) The explicatum should be as *simple* as possible; this means as simple as the more important requirements (1), (2) and (3) permit.

Out of these four criteria, Carnap gave fruitfulness the most evaluative weight. This shouldn't be surprising given that the primary objective of explication is to generate concepts that are useful for specific tasks. Crucially, conceptual engineers also care about fruitfulness. We want our concepts to do things, and we want them to do those things well. As Cappelen candidly puts it, "we should be engaged in two kinds of activities: (i) investigating whether these concepts are defective and (ii) if defects are found, then ameliorating the defective concepts" (Cappelen 2018, 40). So, what work is there left to do? Conceptual engineers care about generating fruitful concepts, and since "fruitfulness is the principal aim when undertaking the method of explication" (Pinder 2020, 2), shouldn't all conceptual engineering employ that method? Curiously, most conceptual engineers do not see it that way. They seem to think that the projects they're interested in fall outside the purview of Explication:

> It's worth noting that explication, as Carnap describes it, is primarily tailored towards improvements appropriate to the languages or conceptual schemes of the 'exact' sciences — physics, mathematics, logic, and the like. By contrast, 'conceptual engineering' (at least, as I'll use the term) covers any form of conceptual improvement. (Nado 2019, 3)

> Carnap aimed to replace certain everyday concepts with 'exact and fruitful' concepts for use in the formal or empirical sciences . . . Carnap was interested in conceptual engineering, primarily in the sense of devising new, technical languages. (Thomasson 2019, 6)
>
> Carnap's notion of explication, however, is narrower than the activity I'm interested in. (Cappelen 2018, 11).
>
> For some purposes, formulating either empirical laws or logical theorems is not what really matters — just think of typical philosophical inquiries, e.g., into the nature of justice or knowledge. (Koch 2019, 702)
>
> Carnap's proposal is not, as it stands, suitable for this particular task. (Pinder 2020, 4)

This reaction is puzzling given that Carnap is explicit about explication being a generalizable methodology, "The task of explication is of very general importance for the construction of concepts." (Carnap 1950, 2)[20] Perhaps the reason why it's presumed explication cannot accommodate other CE projects is that Carnap focused on explicating scientific and logico-mathematical concepts throughout his career. Consequently, we don't have any examples from Carnap clearly demonstrating what explications of social, political, or ethical concepts look like.

The ultimate goal of this chapter is to provide what Carnap did not. I will explain how to extend the method of explication to concepts that are not logical or empirical. In order to do this, first, I will critique a recent account of fruitfulness given by Mark Pinder (2020). Pinder presents and defends what he calls the 'Relevant-Goals Account' of fruitfulness, which he claims "can serve as the aim for conceptual engineering in inquiry in general" (Pinder 2020, 4). I summarize Pinder's view in section two and explain why it fails to deliver the sought-after result. In section three, I present an alternative—pluralist—account of fruitfulness, which I argue makes up for the deficiencies of Pinder's approach. In section four, I describe a case study about gender that illustrates how to effectively use explication for socio-political concepts and reinforces the pluralist notion of

[20] Moreover, Carnap did not see Explication as carving out a distinction between philosophical and scientific concepts, "I see here no sharp boundary line but a continuous transition." (Carnap 1963, 934)

fruitfulness. Section five concludes by drawing out the methodological implications this analysis has for CE.

4.2 The Relevant-Goals Account

Pinder (2020) proposes an account of fruitfulness that he claims overcome the supposed inadequacies of Carnap's view. To start, remember Carnap characterized fruitfulness as "useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept)." (1950, 7) Pinder separates the content of this description into two independent claims: (C1) A concept is fruitful if it furthers the aims of empirical and logical inquiry and (C2) The aim of empirical and logical inquiry is to formulate and confirm universal statements (empirical laws and logical theorems). Pinder then states that:

> Now, the goal in the present paper is to find an account of fruitfulness that can serve as the aim for conceptual engineering in inquiry in general. However, for two reasons, Carnap's proposal is not, as it stands, suitable for this particular task. First, (C1) focuses only on empirical and logical inquiry. But, as illustrated by much of philosophy, inquiry need be neither empirical nor logical—some inquiry is, for example, conceptual or normative. Second, contrary to (C2), the aims of some areas of empirical inquiry are not best understood in terms of universal statements. (Pinder 2020,4)

This quote clearly indicates Pinder belongs to the camp that thinks Explication can't be utilized for CE in general. To this end, he proposes the 'Relevant-Goals Account of Fruitfulness,' which has two components (2020, 6):

> (1) An explicatum is fruitful insofar as its replacement of the corresponding explicandum would facilitate, through the ordinary course of inquiry, progress towards achieving relevant theoretical goals.
>
> (2) Given multiple fruitful candidate explicata for a single explicatum, the most fruitful candidate is that whose replacement of the explicandum would facilitate, through the ordinary course of inquiry, most progress towards achieving relevant theoretical goals.

Of particular importance are the notions of 'theoretical goals' and 'progress.' However, for several reasons, Pinder's account of each falls short. Here is the basic idea, according to Pinder:

> As a first gloss, one might be tempted to characterise theoretical goals simply as goals pertaining to some or other theoretical matter. However, herein, I have a thicker notion in mind. When performing explications, theorists should not aim simply to achieve goals that pertain to some or other theoretical matter. In addition, theorists should aim to thereby make appropriate progress. Now, what constitutes progress in inquiry plausibly varies between different fields: progress in ethics may look nothing like progress in astronomy. To keep things manageable, I will focus on theoretical goals in science. (ibid., 8)

There are three problems with this view. First, even if Pinder's notions of theoretical goals and progress in science were correct, the upshot of his account was supposed to demonstrate how to conceptualize fruitfulness for non-scientific CE projects. The case study Pinder presents to support his account is about the development of a taxonomy for celestial objects, and specifically, definitions of a planet. Pinder claims the account is generalizable, but there was no rationale or concrete examples given to support such an extrapolation. At one point, Pinder even admits features of his account "may not generalise to non-scientific fields of inquiry" (ibid., 10). Thus, it appears that the Relevant-Goals Account of fruitfulness does not extend beyond Carnap's version of Explication at all.

Second, the broad focus on theoretical goals is off-target if the aim is a general account. As mentioned in the introduction, the goals of conceptual engineers are not always theoretical. Thus, any account that aims at generalizing to all types of CE projects needs to include the non-theoretical aspects and goals of conceptual engineers. Furthermore, the goals of scientific inquiry are not always theoretical. Pinder suggests that we understand theoretical goals in science in terms of how they promote theoretical values, what he calls 'value-directedness' (ibid., 9): "A goal is theoretical with respect to T insofar as: ceteris paribus, achievement of that goal would ipso facto transform T into a theory that is better with respect to some theoretical value(s) than T."

Pinder cites Kuhn's well-known list, including internal consistency, coherence, evidential accuracy, scope, explanatory power, and simplicity, as paradigmatic examples of theoretical values in science (Kuhn 1977). He also asserts that the more theoretical values a theory has and the higher

degree to which it has them, the better the theory. This paints an inaccurate picture of the rich and controversial philosophy of science literature on the topic. Pinder acknowledges that "there are also important questions about what the theoretical values in fact are, how they are related, and why they are valued. For present purposes, however, I put such complications and questions aside." (ibid. 9) But, these complications and questions cannot be set aside if the goal is to develop an account of fruitfulness that is inspired by actual scientific practice. The idea that there is a unique and universal set of theoretical values that defines standards of theory assessment for the scientific community arose out of concerns for the underdetermination of theory by evidence, the incommensurability of different paradigms, and the value-ladenness of observations (Duhem 1954; Kuhn 1962; Feyerabend 1962). But, as numerous analyses show, there is not a neat distinction between cognitive or epistemic and so-called 'contextual' values.[21] Relatedly, there are compelling arguments for the inclusion of other values like diffusion of power, novelty, mutuality of interactions, and applicability to human needs in the goals and evaluations of scientific inquiry (Longino 1996). This relates to the final problem with Pinder's account, which is his notion of scientific progress.

Like the controversy surrounding theoretical values in science, the concept of scientific progress has a long and thorny history in the philosophy of science. There are three traditional approaches to characterizing scientific progress. First is an epistemic approach that measures progress in terms of acquiring knowledge (Bird 2007). A second, semantic approach, gauges progress in terms of scientific theories being nearer to the truth (Niiniluoto 2014). Finally, there is a functional-internalist approach that measures progress in terms of 'problem-solving' and 'problem defining' (Kuhn 1962; Laudan 1977; Shan 2019). Importantly, none of these approaches account for

[21] Contextual values (also sometimes called 'non-cognitive' values) are moral, personal, social, political, and cultural values like pleasure, justice, equality, conservation of the environment, and diversity. Contextual values supposedly conflict with the 'cognitive' or 'epistemic' values typically used to characterize scientific reasoning and are usually thought of as truth conducive. See chapters four and five in Longino 1990.

progress in applied ethically-driven sciences like conservation biology. As Justus and Wakil argue, "These sciences don't deliver anything resembling justified true beliefs about a mind-independent cosmos. Instead, they supply data-driven, algorithmically rigorous, evidence-based means for achieving ethical goals". Justus and Wakil's analysis supports a fourth, alternative, concept of scientific progress. Presented by Heather Douglas, her view does away with the applied/theoretical dichotomy and instead defines scientific progress "in terms of the capacity to predict, control, manipulate, and intervene in various contexts" (Douglas 2014, 62).

With this very brief background of the debate, it's clear how Pinder's narrow construal of scientific progress has obscured his analysis:

> As it stands, Value-Directedness may not generalise to non-scientific fields of inquiry. Several questions are relevant here. What exactly are scientific theories? To what extent are they like the theories we find in other fields of inquiry? Are theoretical values applicable to non-scientific theories and, if so, to what extent? Does progress in other disciplines consist of developing better theories? If we think of philosophy as broadly continuous with science, then we might think that the traditional theoretical values apply mutatis mutandis in philosophy. *But, in contrast, if we focus instead on the normative nature of philosophical inquiry, then many of the traditional theoretical values seem inappropriate.* It is unclear that we can straightforwardly measure progress in ethics with reference to evidential accuracy or explanatory power, for example. These are deep issues about the nature of scientific and other inquiry, and I cannot resolve them here. For present purposes, I settle for Value-Directedness, construed only as an account of theoretical goals in scientific inquiry. (Pinder 2020, 10; italics my emphasis)

The italicized portion above suggests that Pinder seems to think that science and scientific progress, in particular, is not normative. Such a view coincides closest with the epistemic account of scientific progress. But as already noted, this is only one of multiple concepts of scientific progress, and it is one that has been heavily criticized in recent years. Taking into consideration the complexities of defining theoretical goals and scientific progress reveals how some of Pinder's questions are misinformed. I want to focus on two of them:

1. **"Does progress in other disciplines consist of developing better theories?"**

   Answer: First, progress in science consists of more than developing better theories. If

we are going to make a comparison across disciplines (also note that discipline boundaries can be fuzzy). It's important we compare the various kinds of progress each field makes. Second, once it's explicit how progress is being defined, then progress in science can look like progress in philosophy. For example, someone who endorses the semantic approach will say that progress has been made when a scientific theory gets us closer to the empirical truths. Similarly, a robust moral realist is going to judge an ethical theory that gets closer to the moral truth as progress. Pragmatists will likely judge ethical progress along similar lines as the truth-functional internalist about scientific progress. And progress for philosophers with explicitly practical goals will look very similar to the capacities to predict, control, manipulate, and intervene laid out in Douglas's account of scientific progress.

2. **"Are theoretical values applicable to non-scientific theories and, if so, to what extent?"**

   Answer: If we take seriously the criticism that the traditional theoretical values are not reliably truth conducive, and we agree that contextual values should play a role in scientific practice, then we have reason to think that at least some of these values will be applicable in non-scientific situations. Here is a non-exhaustive list of proposed values from the philosophy of science literature: consistency, coherence, evidential accuracy, novelty, mutuality of interaction, scope, explanatory power, simplicity, diffusion of power, unification, and applicability to human needs. It should be obvious that some of the values on this list pertain to philosophy as well. Ethicists, metaphysicians, logicians, etc., care about their theories being consistent, novel, and unifying, for example. Pinder further claims that "it is unclear that we can straightforwardly measure progress in ethics with reference to evidential accuracy or explanatory power." But the answer does become clear once we specify what the goal of ethics is. Just like science has a variety of goals that are not always

theoretical, ethics also has a variety of non-theoretical goals. The case study in section four will demonstrate how evidential accuracy and explanatory power are important values for ethics, social, and political philosophy.

The take away from this section is the following. First, scientific practice is messy and complicated—so is philosophy. Any comparison between the two needs to account for the complexities of both domains. Pinder's naive gloss on the philosophy of science reveals the shortcomings of failing to do this. Second, I take the variety of concepts regarding the theoretical value and scientific progress to suggest that any general account of fruitfulness is misguided. Why would we want or expect a general account of fruitfulness (or any of the explication success conditions) when the goals of explicators vary so widely? A generalist account cannot capture the diversity of goals and practices in science or philosophy, and thus is ill-suited to be part of CE's methodology. In the next section, I propose a pluralist account of fruitfulness that can.

4.3 A Pluralist Account of Fruitfulness

In this section, I propose a pluralist account of fruitfulness that coheres with a wide range of CE projects and goals. Before presenting the view, it's worth noting that Pinder anticipated the need for more specificity with regards to fruitfulness than the Relevant-Goals account provides. So, he supplements his account with another view he calls *Contextualism,* which is supposed to specify which theoretical goals are relevant in a given case:

> Contextualism is intended to capture this thought by tying relevant theoretical goals to whoever or whatever is performing a given explication. More precisely, Contextualism is the view that, for any given explication, the relevant theoretical goals are the explicator's theoretical goals . . . On this view, explicators have a significant degree of control over which of their theoretical goals are relevant, and thus over the measure of fruitfulness that applies in a given case. This does not imply that anything goes. . . as relevant theoretical goals are always theoretical goals, they must (by Value-Directedness) be directed towards the improvement of certain theories in certain respects . . . The Relevant-Goals Account is thus general enough to apply to any given explication. And, by providing additional details about the explicator's theoretical goals, we can derive a substantive measure for fruitfulness in that case. (Pinder 2020, 10)

Contextualism is on the right track but does not go far enough for two reasons. First, as the last section noted, philosophers and scientists have a variety of goals, many of which are not theoretical. Pinder's concern not to have a measure of fruitfulness collapse into total relativism is laudable. But his criteria of Value-Directedness only pertains to improving theories. Thus, as currently formulated, it does not help derive non-theoretical measures of fruitfulness. Second, contextualism broadly construed is a doctrine that, as the name suggests, emphasizes the importance of the context of inquiry. The basic idea is that evaluations of actions, utterances, and expressions can only be understood relative to a particular context. Pluralism, on the other hand, is a doctrine that maintains that two or more states, groups, frameworks, principles, projects, sources of authority, etc., can coexist. In other words, pluralism recognizes many goals and evaluative standards as equally valuable.[22] Indeed, one thing that makes CE such an attractive and interesting research program is the wide range of projects emerging from it. Contextualism does not secure the value of different CE projects, whether those projects are scientific or not and whether the goals are theoretical or not. So, in order to uphold and respect the diversity of CE, we need a pluralist account of fruitfulness.[23] Here is a plain statement of the account:

(1) Definition of progress: "if the aim of X is Y, then X makes progress when X achieves Y or promotes the achievement of Y." (Bird 2007, 83)

(2) When Y = [an explicator's goal(s)] and X = [a proposed explicatum E], E is fruitful in so far as it makes progress as defined by (1).

[22] Notably, pluralism is at the heart of Carnap's famous Principle of Tolerance: "It is not our business to set up prohibitions, but to arrive at conventions." (Carnap 1937, 51)

[23] Other endorsements of pluralism with respect to CE can be found in Burgess and Plunkett (2013) and Brigandt and Rosario (2020)

(3) Given multiple fruitful explicata $E^{(1...n)}$, E_m is considered the most fruitful to the degree it achieves or promotes the achievement of Y more than the alternatives.

(4) Explicata may be fruitful in distinct ways. Thus, there may not be one explicatum that is the most fruitful.

(5) Given multiple explicanda, we must clearly distinguish between an evaluation of the explicatum and the explicandum.

Further remarks about these components are in order. To start, (1) is a definition of progress that is not overly narrow. It does not construe progress as merely theoretical, as Pinder's account does. And it is applicable to any domain of inquiry, not just science. As such, Pinder's comparative concerns regarding what progress looks like in non-scientific fields dissolves. Components (2) and (3) characterize and gauge fruitfulness by its progressive features as defined by (1). Again, this avoids the theoretical trappings of Pinder's original proposal, but it maintains the importance of the explicator's goals in deriving measures of fruitfulness. Finally, components (4) and (5) capture the pluralism we want with respect to different CE projects. There are potentially several different and non-equivalent ways an explicatum can achieve or promote the achievement of an explicator's goal.[24] Accordingly, component (4) appropriately constrains (3) by permitting the acceptance of different explicata that are fruitful in distinct ways. Component (5) is an important methodological reminder. Because an explicatum's fruitfulness is gauged relative to the specified goal, and because a single term can pick out different concepts, we need to ensure that our evaluations of fruitfulness pertain exclusively to the intended goal. Failure to do so results in philosophers talking past each other or engaging in what Chalmers calls 'merely verbal disputes' (Chalmers 2011). This was, in fact, Carnap's great insight into the problem of probability:

[24] Justus (2012), for example, demonstrates how different explications of 'ecological stability' are fruitful but in distinct ways.

> When we criticize the theory of probability proposed by an author, we must clearly distinguish between a rejection of his explicatum and a rejection of his explicandunum . . . Most investigators in the field of probability apparently believe that all the various theories of probability are intended to solve the same problem and hence that any two theories which differ fundamentally from one another are incompatible . . . This is primarily due to the unfortunate fact that both concepts are designated by the same familiar, but ambiguous word 'probability' . . . This whole controversy seems to me futile and unnecessary. The two sides start from different explicanda, and both are right in maintaining the scientific importance of the concepts chosen by them as explicanda. (Carnap 1945, 518-520)

The spirit of Carnap's analysis should extend to all CE projects. For any given concept, there are likely several different things a philosopher may want the concept to do. Just like there are different equally important explicanda for the concept of probability, there will be equally important explicanda for a myriad of concepts across different areas of philosophy. Accordingly, we need to make sure our evaluations of an explicatum's fruitfulness is gauged with respect to the appropriate explicandum. One way to help ensure the five components of this account are met would be to follow a clear step by step procedure for doing explications. The next section provides a roadmap for undertaking explications and illustrates how to use it through an example of explicating gender concepts.

4.4 How to Explicate Gender Concepts (and anything else)

The goal of this section is two-fold. First, it provides an easy to follow roadmap for doing explications that any conceptual engineer can use regardless of the kind of concept they are trying to engineer. Second, it provides an example of engineering gender concepts to demonstrate the roadmap's utility and provides the long sought after case of explicating non-scientific concepts. The reader should refer to figure 4.1 to accompany the description of the procedure.

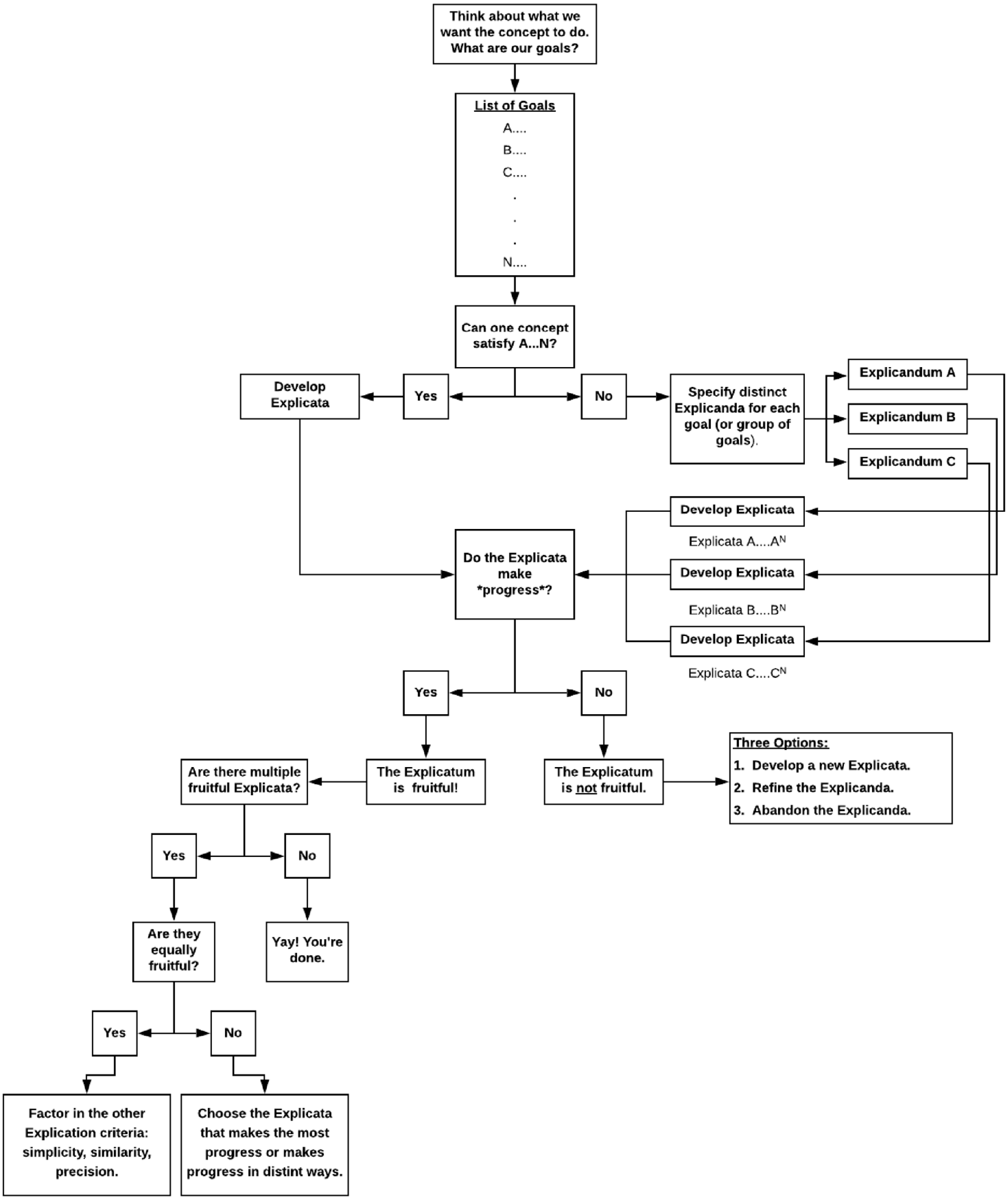

Figure 4.1: Explication Flow Chart.

The first step in doing an explication is to ask what work we want the concept to do for us and make a list of those goals. Looking at the list, an explicator then decides whether it's likely one concept will be able to satisfy every goal. Some relevant factors to consider here would be empirical evidence about the concept's current use, previous attempts at explicating the concept, whether the goals differ in kind (practical vs. theoretical), and any seeming inconsistencies between goals. If the explicator determines that one concept might be sufficient to satisfy every goal, then they begin developing the corresponding explicatum (more on this below). If, however, it seems unlikely that one explicatum can capture the entire list, then the explicator should specify multiple explicanda. There could be one explicandum per goal, or the explicanda could be defined by grouping various goals together. These groupings are up to the explicator's discretion. In making such determinations, the factors mentioned above are also applicable.

Once the explicandum or explicanda has been specified, the next task is to develop the explicata. Creating an explicatum is typically a very difficult task. It requires a great deal of cognitive creativity and effort to develop clear and explicit conceptual definitions to satisfy particular goals. (Thagard 1992; Sinatra and Pintrich 2003) Importantly, in developing explicata, it is crucial to take the other explication success conditions—simplicity, similarity, and precision—into consideration. Though fruitfulness is given the most evaluative weight, the significance of the other conditions will likely vary depending on the kind of concept being explicated. It's undeniable that scientists are not hostage to the similarity criterion. As Justus (2012, 162) puts it, "scientific definitions readily sacrifice intuitive accord with the meaning of the (often problematically vague) concept being defined to enhance formal rigor, experimental testability, measurability in the field, mathematical and theoretical unification, etc." But for non-scientific concepts, it may be appropriate to give similarity non-trivial weight. As we will see in the case study below, successfully explicating socio-political concepts will sometimes require prioritizing similarity for specific normative goals. Likewise,

simplicity could matter a great deal for certain practical agendas. If an explicata is too difficult to understand or implement, then it won't make progress as defined by (1) and thus be deemed unfruitful.

Finally, although subsidiary to fruitfulness, precision is paramount for all CE projects for two reasons: First, precision is instrumentally valuable for fruitfulness as increases in precision typically accompany increases in fruitfulness. Second, one of the ways precision increases fruitfulness is by increasing measurability/testability (Justus 2012). This is important for evaluating the success of our conceptual revisions most generally, which is the degree to which an explicated concept achieves its intended function better than the original. Ultimately evaluating that success requires data about the function of the original concept and data about the function of the explicated concept. Insofar as precision increases measurability, developing more precise explicata will generally be of great utility for CE. (See chapter 1 of this dissertation.) Again, defining explicata is a difficult intellectual project, but taking the other explication criteria into consideration from the start can help inform and guide the development of the respective explicata.

After an explicatum has been introduced, the pertinent question is whether it's fruitful. Recall, on the pluralist account, fruitfulness is characterized in terms of progress: if the goal of the explicatum is X, then the explicatum makes progress when it achieves X or promotes the achievement of X. When an explicatum makes progress toward the relevant goal it is fruitful, and how fruitful the explicatum is will be determined by how much progress it makes. If an explicatum is not fruitful, an explicator has three options: (1) try to improve or develop a new explicatum, (2) go back and refine the explicanda, or (3) abandon the concept entirely because it just can't be improved. Which option an explicator takes will depend on why the proposed explicata is not fruitful and psychological facts about themselves: which option do they *want* to take? How many times have they tried explicating the concept? Can they think of any new explicata? Etc.

If there are several fruitful explicatums, the explicator should look at how many fruitful explicata per explicanda there are. If there is only one fruitful explicatum, then the explication process is complete. However, if there are multiple fruitful explicata, the explicator should evaluate whether the explicata are equally fruitful—i.e., do they make similar degrees of progress? The answer could be no for two reasons: First, there could be one explicata that is more fruitful than the rest; in which case this is the explicata to choose. Second, as mentioned before, explicata can make progress (and thus be fruitful) in distinct ways. In this scenario, it is appropriate to accept different explicata that are fruitful for certain purposes in certain contexts. On the other hand, if the explicata are fruitful in non-distinct ways, then the other explication criteria should be brought to bear on our conceptual decision-making. Again, the relative weight and potential trade-offs between the other criteria are made on a case by case basis. This kind of conceptual flexibility is a benefit of the pluralist approach, as different kinds of concepts may require prioritizing different criteria. With this flowchart and general description of the procedure laid out, next, I will present a case study that demonstrates the method and reinforces the pluralist account of fruitfulness.

Without even mentioning Carnapian explication, a recent CE analysis about gender concepts effectively illustrates the utility of the explicative methodology. Brigandt and Rosario (2020) provide a novel and significant contribution to the discussion on how to conceptualize race and gender. For the purposes of this paper, I will focus only on their accounts of gender. Brigandt and Rosario begin by acknowledging there are several things philosophers may want the concept of gender to do, what the flowchart calls the list of goals:

A. Identify and explain gender-based discrimination.

B. Assign legal rights and ensure gender-appropriate social recognition.

C. Empower persons by means of their gender identity.

D. Be trans-inclusive.

The next step is to ask whether one concept of gender could satisfy A-D. Brigandt and Rosario respond in the negative: "Our general position will be that there is no unique, privileged concept of gender or race. Rather than an alleged all-purpose concept, what is needed is a plurality of concepts of gender and race, each of which is geared towards certain epistemic and/or social aims" (Brigandt and Rosario 2020, 100). The next step is to specify distinct explicanda for these functions. Brigandt and Rosario state that for the descriptive purposes of identifying and explaining gender-based discrimination, being trans-inclusive is not a necessary desideratum: "One reason is that relative to one explanatory aim, a concept should only include what is explanatorily relevant and thus not adduce any feature (no matter how true) that is irrelevant to the phenomena to be explained. Therefore, any explanatory concept must exclude some aspects of reality, which however have to be included with respect to some other explanatory aim" (ibid., 115). That said, for the other functions, (B) Assign legal rights and ensure gender-appropriate social recognition and (C) Empower persons by means of their gender identity, the inclusion of trans people is a necessary condition for a successful explicata. Thus, they propose three gender explicanda, which we will label like the following:

Gender A: Identify and explain gender-based discrimination.

Gender BD: Assign legal rights and ensure gender-appropriate social recognition for trans and non-trans persons.

Gender CD: Empower trans and non-trans persons by means of their gender identity.

Next, they begin the process of developing explicatums for the respective explicandums. In doing so, they consider many of the factors mentioned above, like previous proposals, evidence about the concept's current use, whether similarity to the original concept matters, etc. Below is a summary of their theorizing about the development of each explicatum.

Gender A:

Brigandt and Rosario start by analyzing Sally Haslanger's account of 'woman.' That account provided the first definitional concept of women in the philosophical literature; and importantly rebutted previous skepticism that a coherent concept of gender was even possible.

S *is a woman* if

> (i) S is regularly and for the most part observed or imagined to have certain bodily features presumed to be evidence of a female's biological role in reproduction;
>
> (ii) that S has these features marks S within the dominant ideology of S's society as someone who ought to occupy certain kinds of social position that are in fact subordinate (and so motivates and justifies S's occupying such a position); and
>
> (iii) the fact that S satisfies (i) and (ii) plays a role in S's systematic subordination, that is, *along some dimension*, S's social position is oppressive, and S's satisfying (i) and (ii) plays a role in that dimension of subordination. (Haslanger 2000, 42)

There are a few positive features of this initial definition Brigandt and Rosario identify. First, it does not assume that women must have a particular shared psychological identity or identical social experiences. This is crucial as considerations of intersectionality clearly demonstrate that different women have varying social experiences due to occupying different social positions and facing different social expectations. Second, contrary to other accounts, it does not construe 'women' as a purely biological category. Instead, Haslanger's definition rightly treats women as a biosocial kind: phenomena that are mutually causally influenced by biological and social factors. Given the stated goal of identifying and explaining persistent patterns of unjust inequalities between males and females, a successful explication of 'woman' must include the relevant causal factors and how they interact. By explicitly including social factors, Haslanger's definition succeeds in this regard.

That said, Brigandt and Rosario highlight various ways in which Haslanger's original definition might be improved. First, they suggest rather than providing ameliorative definitions of 'woman' and 'man'; the focus should be on 'gender' specifically. This supposedly has two benefits.

First, the focus is not on the extension of a concept, which is the target of several criticisms of Haslanger's definitions. For instance, Saul 2006 argues that Haslanger's account problematically defines women as female persons who are subordinated. This forecloses the question of whether there are women that are not subordinated. She also notes that if there is a woman who has not faced discrimination, or a woman who overcomes discrimination, they will not count as women on Haslanger's definition. Sometimes the specific extension of a concept is relevant (e.g., for the purpose of assigning legal rights). But, Brigandt and Rosario argue that for the purposes of identifying discrimination, any account of gender: "[S]hould be open to the boundaries of different genders being vague, to some persons being of more than one gender, and to there being more than two genders. An account of 'gender' is definitely not committed to any of this (while substantial expectations may obtain for a proposed definition of 'woman' and 'man'), and can even make room for the number of different genders varying across cultural history" (ibid., 116).[25]

Second, Haslanger advocates for a society where there are no subordinated women (as defined by her). But this has been problematically viewed by some as committing to the goal of 'the elimination of women.' Such an understanding "may well alienate many of those who would be most likely to support anti-oppression movements." (Saul 2006, 139). Brigandt and Rosario claim their account circumvents this problem, as the agenda of eliminating systemic injustice and subordination can be framed much more palpably in terms of the elimination of 'gender' rather than the elimination of 'women.'

In addition to the specific focus on gender, Brigandt and Rosario argue that the concept should be relational "in which the construal of one gender makes reference to other genders or to the properties used to characterize other genders." (2020, 116). Haslanger's account and other

[25] Notice how this consideration tacitly pertains to the precision criterion. In particular, that some vagueness (in this specific case) is warranted given the stated goals.

prominent alternatives (Saul 2006; Jenkins 2016), all attempt to articulate intrinsic features of the concept. Merely appealing to features that are supposedly intrinsic makes it difficult to capture any variation within the conceptual category. In contrast, a relational account can identify the *differential* social treatment of genders:

> It neither has to claim that all women are subordinated — as Haslanger's definition problematically does by excluding non-subordinated females — nor does it have to articulate what this allegedly shared social subordination consist of, e.g., the 'certain kinds of social position' to which Haslanger appeals in condition (ii) of her definition… Instead, a relational construal can indicate that compared to men, women tend to be discriminated against because of their gender. (Brigandt and Rosario 2020, 116)

Finally, a relational account of 'gender' rather than 'woman' coheres with considerations of intersectionality. It only identifies gender as one possible axis of discrimination and allows for other social identities (race, class, sexual orientation, trans status, etc.) as other dimensions of discrimination which intersect differently for individuals such that they experience different kinds, and degrees of, oppression and privilege (Crenshaw 1991).

Gender BD:

The goal of this explicanda is to assign legal rights and ensure gender-appropriate social recognition for trans and non-trans persons. Brigandt and Rosario begin by noting how their proposals for gender A won't suffice:

> [A] relational concept of gender fails to delineate exactly which persons are of a certain gender, which is not relevant for the purpose of identifying and explaining society-wide discrimination, but is indeed indispensable for the aim of ensuring gender-based legal rights and social recognition in line with someone's gender identity. With respect to this aim, a different concept of gender has to be employed, that entails a specific extension of 'woman' and other gender categories. (ibid., 117)

They first consider Jenkin's (2016) account, which was developed specifically to be trans-inclusive. Jenkins makes a distinction between gender as class (essentially Haslanger's concept) and gender as an identity, which is what her account attempts to capture: "S has a female gender identity iff S's

internal 'map' is formed to guide someone classed as a woman through the social or material realities that are, in that context, characteristic of women as a class" — (Jenkins 2016, 410).

Jenkins claims that this definition ensures that all trans women count as having a female gender identity. But Brigandt and Rosario reveal why that isn't correct. Jenkins's notion of an 'internal map' requires that a person is aware of gender-specific social norms and expectations. For example, for women, "Her experience of social and material reality includes navigating the norm that women should have hairless legs, even though she is not complying with it" (Jenkins 2016, 411) Brigandt and Rosario rightly point out that this potentially excludes not only trans women but any women who, for whatever reason, are not aware of the norm that they should shave their legs. Given the stated goal, Brigandt and Rosario conclude that no successful definition for Gender BD will make a person's gender contingent on any kind of ideological test, like awareness of gender-related norms. In theorizing about an alternative to Jenkins account, Brigandt and Rosario examine an interesting analogy to political party membership:

> Republicans tend to have many factual beliefs, political values, and social practices in common that form their political identity, and these shared features are important to sociologists and political scientists for predictive and explanatory purposes (predictions and explanations that are useful even if there are some Republicans that do not conform to the stereotype). However, Republicans who are untypical in terms of their beliefs and values can still be party members and have concomitant rights. For instance, being registered as a Republican voter (according to state guidelines) suffices for being eligible to vote in a Republican primary — no specific ideological test is required to possess this voting right. (Brigandt and Rosario 2020, 112)

Just as the only requirement to be recognized and guaranteed certain rights as a Republican is being registered as such, receiving gender-appropriate social recognition and rights likewise should only require self-identification with the gender category. Brigandt and Rosario conclude, "we advocate using a concept of gender that defines someone to be a 'woman' or 'man' solely in terms of this person's sincere self-identification with a particular gender, for example, as long as someone identifies as belonging to the gender commonly denoted by the term 'woman'" (ibid., 113).

Gender CD:

This last explicandum has the social aim of empowering all persons by means of their gender identity. Gender A is intended to be useful for identifying and advancing our understanding of socially influenced gender-based discrimination, which in turn may inform public policy initiatives for reducing discrimination. But as Brigandt and Rosario note, this might not help individual women empower themselves in their everyday lives. Achieving that goal seems to require positive psychological aspects (as opposed to internalized harmful stereotypes) that are conducive to feeling personally empowered. Gender BD purposely omitted such psychological features and therefore is unsuitable for the goal of personal empowerment.

Given the individualistic nature of feeling empowered, it is unlikely that a single explicata can be given for an identity concept like Gender CD. Accordingly, Brigandt and Rosario do not specify any specific explicata for Gender CD. However, they suggest that "[a] person should be able to pick and choose amongst a variety of (nonbinary) gender-associated behaviours and activities in order to construct one's own particular gender attributes that enable agency" (ibid., 118). Although they did not propose specific explicata in their own analysis, the fact that multiple different Gender CD explicata will be needed for empowering different groups of persons and individuals is something Explication, and the pluralist account of fruitfulness, in particular, welcomes and encourages.

4.5 Conclusion

I want to conclude with some general methodological remarks about using Explication for the purposes of conceptual engineering. Two recent accounts of explication have been proposed by Cordez (2020) and De Benedetto (2020). Cordez's 'structural account' provides a step-by-step guide for explications but deliberately excludes Carnap's four requirements (similarity, exactness, fruitfulness, simplicity) and consists of 19 formal postulates and 39 definitions in a first-order predicate language with standard set theory. De Benedetto similarly 'explicates explication' using

conceptual spaces theory and particularly focuses on establishing relations between explicanda and explicata using geometrical or topological dimensions. Both these accounts are novel and technically sophisticated but come at the cost of general practicality. Perhaps such accounts will prove effective for formalized conceptual engineering projects. However, as pointed out several times now, the goals of conceptual engineers vary widely (especially for socio-political projects). Such variation establishes a need for a simpler and easy to follow methodology that does not depend on formalized techniques. The pluralist account of fruitfulness presented here and accompanying road map was created to provide just that.

Finally, acknowledging that different methodologies may be useful for different kinds of conceptual engineering projects is, notably, a conclusion in line with Carnap's pluralist and cooperative attitude: "We all agree that it is important that good analytic work on philosophical problems be performed. Everyone may do this according to the method, which seems to be the most promising to him. The future will show which of the two methods, or which of the many varieties of each, or which combinations of both, furnishes the best results" (Carnap 1963, 939).

As stated in the introductory chapter, the thesis of this dissertation is that Carnapian Explication is the methodology that conceptual engineers should be following. In chapter two, I argued that empirical investigation is required to vet the success of CE proposals, and thus the precision criterion plays an indispensable role in generating concepts that are conducive to those empirical methods. In chapter three, I used those empirical methods to support a recommendation of a wholesale abandonment of the parsimony concept. That chapter also analyzed a case study that demonstrated the importance of the explicanda and explicata distinction. This distinction was incorporated into the pluralist account of fruitfulness I presented in the final chapter. The pluralist account is what undergirds the ultimate claim made here: that Explication is applicable to any concept, not just logical or empirical. The significance of this cannot be stated enough as this is the

primary contention against Explication (even by neo-Carnapian conceptual engineers). By providing an easy-to-follow flowchart and using it to track a conceptual engineering analysis of gender concepts, I hope the pervasive idea that Explication is inadequate for the general purposes of conceptual engineering is finally laid to rest.

BIBLIOGRAPHY

Adleberg, Toni, Morgan Thompson, and Eddy Nahmias (2015). "Do Men and Women Have Different Philosophical Intuitions? Further Data". In: *Philosophical Psychology* 28.5, pp. 615–641. DOI: 10.1080/09515089.2013.878834.

Baker, A. (2007). "Occam's Razor in Science: A Case Study From Biogeography." In: Biology and Philosophy 22.2, pp. 193–215. DOI: 10.1007/s10539-006-9027-9.

Bird, Alexander (2007). "What is Scientific Progress?" In: Noûs 41.1, pp. 64–89. DOI: 10.1111/j.1468-0068.2007.00638.x.

Blanchard, Thomas, Tania Lombrozo, and Shaun Nichols (2018). "Bayesian Occam's Razor Is a Razor of the People." In: *Cognitive Science* 42.4, pp. 1345–1359. DOI: 10.1111/cogs.12573.

Brigandt, Ingo, and Esther Rosario (2020). "Strategic Conceptual Engineering for Epistemic and Social Aims." In: *Conceptual Engineering and Conceptual Ethics*. Ed. by Alexis Burgess, Herman Cappelen, and David Plunkett. Oxford: Oxford University Press, pp. 100–124.

Brun, Georg (2016). "Explication as a Method of Conceptual Re-Engineering." In: *Erkenntnis* 81.6, pp. 1211–1241. DOI: 10.1007/s10670-015-9791-5.

Burgess, Alexis, and David Plunkett (2013). "Conceptual Ethics I." In: *Philosophy Compass* 8.12, pp. 1091–1101. DOI: 10.1111/phc3.12086.

Cappelen, Herman (2018). *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

Carnap, Rudolf (1945). "The Two Concepts of Probability: The Problem of Probability." In: *Philosophy and Phenomenological Research* 5.4, pp. 513–532. DOI: 10.2307/2102817.

– (1950). *Logical Foundations of Probability*. University of Chicago Press.

– (1963a). "Intellectual Autobiography." In: *The Philosophy of Rudolf Carnap*. Ed. by Rudolf Carnap. Open Court, pp. 3–84.

– (1963b). "Replies and Systematic Expositions." In: *¸Iteschilpp: Prc*. Ed. by Paul Arthur Schilpp. Open Court, pp. 859–1013.

Carus, A. W. (2007). *Carnap and Twentieth-Century Thought: Explication as Enlightenment*. Cambridge University Press.

Chalmers, David J. (2011). "Verbal Disputes." In: *Philosophical Review* 120.4, pp. 515– 566. doi: 10.1215/00318108-1334478.

Coelho, Marco TÃºlio P., JosÃ c Alexandre Diniz-Filho, and Thiago F. Rangel (2019). "A parsimonious view of the parsimony principle in ecology and evolution." In: *Ecography* 42.5, pp. 968–976. DOI: 10.1111/ecog.04228.

Colaço, David et al. (2014). "Epistemic Intuitions in Fake-Barn Thought Experiments." In: *Episteme* 11.2, pp. 199–212. DOI: 10.1017/epi.2014.7.

Crenshaw, Kimberle (1991). "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." In: *Stanford Law Review* 43.6, pp. 1241–1299. URL: http://www.jstor.org/stable/1229039.

Cruz, Helen De (2017). "Religious Disagreement: An Empirical Study Among Academic Philosophers." In: *Episteme* 14.1. DOI: 10.1017/epi.2015.50.

Dixson, Alan (2015). "Human sexuality and the menstrual cycle: comment on Havliček et al." In: *Behavioral Ecology* 26.5, pp. 1261–1261. ISSN: 1045-2249. DOI: 10.1093/ beheco/arv085.

Duhem, Pierre (1954). *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.

Eklund, Matti (2002). "Inconsistent Languages." In: *Philosophy and Phenomenological Research* 64.2, pp. 251–275. DOI: 10.1111/j.1933-1592.2002.tb00001.x.

– (2015). "Intuitions, Conceptual Engineering, and Conceptual Fixed Points." In: *The Palgrave Handbook of Philosophical Methods*. Ed. by Christopher Daly.

Eklund, Matti (2020). "Variance Theses in Ontology and Metaethics." In: *Conceptual Engineering and Conceptual Ethics*. Ed. by Alexis Burgess, Herman Cappelen, and David Plunkett.

Feyerabend, Paul K. (1962). "Explanation, Reduction and Empiricism". In: *Crítica: Revista Hispanoamericana de Filosofía*. Ed. by H. Feigl and G. Maxwell, pp. 103–106.

Fraassen, van, (1980). *The Scientific Image*. New York: Oxford University Press.

Frege, Gottlob (1953). *The Foundations of Arithmetic*. Evanston: Ill., Northwestern University Press.

Gangestad, Steven and Nicholas M. Grebe (2015). "Are within-cycle variations in women's sexual interests mere by-products? A comment on Havliček et al." In: *Behavioral Ecology* 26.5, pp. 1262–1263. DOI: 10.1093/beheco/ arv113.

Gildersleeve, Kelly, G Haselton, and Melissa R Fales (Sept. 2014). "Do women's mate preferences change across the ovulatory cycle? A meta-analytic review". In: *Psychological bulletin* 140.5, 1205-1259. ISSN: 0033-2909. DOI: 10.1037/a0035438.

Goodman, Nelson (1955). *Fact, Fiction, and Forecast*. Harvard University Press.

Haselton (2015). "A new (or truly parsimonious) perspective? A comment on Havliček et al." In: *Behavioral Ecology* 26.5, pp. 1263–1264. doi: 10.1093/beheco/arv116. Haselton, Martie and

Kelly Gildersleeve (2011). "Can Men Detect Ovulation?" In: *Current Directions in Psychological Science* 20.2, pp. 87–92. DOI: 10.1177/0963721411402668.

Haslanger, Sally (2000). "Gender and Race: (What) Are They? (What) Do We Want Them to Be?" In: *Noûs* 34.1, pp. 31–55. DOI: 10.1111/0029-4624.00201.

– (, 2012). *Resisting Reality: Social Construction and Social Critique*. Oxford University Press.

Havliček, Jan et al. (2015a). "Greater precision, not parsimony, is the key to testing the peri-ovulation spandrel hypothesis: a response to comments on Havliček et al. 2015". In: *Behavioral Ecology* 26.5, pp. 1265–1267. DOI: 10.1093/beheco/arv137.

– (2015b). "The spandrels of Santa Barbara? A new perspective on the peri ovulation paradigm". In: *Behavioral Ecology* 26.5, pp. 1249–1260. ISSN: 1045-2249. DOI: 10.1093/beheco/arv064.

Holt, Robert D. (2009). "Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives". In: *Proceedings of the National Academy of Sciences* 106.Supplement 2, pp. 19659–19665. DOI: 10.1073/ pnas.0905137106.

Jeffrey, Richard C. (1965). *The Logic of Decision*. University of Chicago Press.

Jeffreys, H. (1961). *Theory of Probability*. Clarendon Press, Oxford.

Jenkins, Katharine (2016). "Amelioration and Inclusion: Gender Identity and the Concept of Woman." In: *Ethics* 126.2, pp.394-42. DOI: 10.1086/683535.

Justus, James (2012). "Carnap on Concept Determination: Methodology for Philosophy of Science." In: *European Journal for Philosophy of Science* 2.2, pp. 161–179. DOI: 10.1007/s13194-011-0027-5.

Knobe, Joshua (2003). "Intentional Action in Folk Psychology: An Experimental Investigation." In: *Philosophical Psychology* 16.2, pp. 309–325. DOI: 10.1080/09515080307771.

Koch, Steffen (2019). "Carnapian Explications, Experimental Philosophy, and Fruitful Concepts." In: *Inquiry: An Interdisciplinary Journal of Philosophy* 62.6, pp. 700– 717. DOI: 10.1080/0020174x.2019.1567381.

Kuhn, Thomas S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press.

Kuhn, Thomas S. (1977). "Objectivity, Value Judgment, and Theory Choice." In: *The Essential Tension: Selected Studies in Scientific Tradition and Change*. University of Chicago Press, pp. 320–39.

Lange, Marc (1995). "Spearman's Principle." In: *British Journal for the Philosophy of Science* 46.4, pp. 503–521. DOI: 10.1093/bjps/46.4.503.

Laudan, L. (1977). *Progress and its Problems: Toward a Theory of Scientific Growth*. University of California Press.

Lindauer, Matthew (2019). "Experimental Philosophy and the Fruitfulness of Normative Concepts." In: *Philosophical Studies*, pp. 1–24. DOI: 10.1007/s11098-019-01302-3.

Lombrozo, Tania (2007). "Simplicity and probability in causal explanation." In: *Cognitive Psychology* 55, pp. 232–257.

Longino, Helen E. (1990). *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton University Press.

– (1996). "Cognitive and Non-Cognitive Values in Science: Rethinking the Dichotomy." In: *Feminism, Science, and the Philosophy of Science*. Ed. by Lynn Hankinson Nelson and Jack Nelson. Kluwer Academic Publishers, pp. 39–58.

Machery, Edouard (2009). *Doing Without Concepts*. Oxford University Press.

– (2017). *Philosophy Within its Proper Bounds*. Oxford University Press.

Machery, Edouard et al. (2004). "Semantics, Cross-Cultural Style." In: *Cognition* 92.3, pp. 1–12. DOI: 10.1016/j.cognition.2003.10.003.

May, Joshua et al. (2010). "Practical Interests, Relevant Alternatives, and Knowledge Attributions: An Empirical Study." In: *Review of Philosophy and Psychology* 1.2, pp. 265–273. DOI: 10.1007/s13164-009-0014-3.

Nado, Jennifer (2015). "Philosophical Expertise and Scientific Expertise." In: *Philosophical Psychology* 28.7, pp. 1026–1044. DOI: 10.1080/09515089.2014.961186.

Nado, Jennifer (2019a). "Conceptual Engineering Via Experimental Philosophy." In:*Tandf: Inquiry*, pp. 1–21.

– (2019b). "Conceptual Engineering, Truth, and Efficacy." In: *Synthese*, pp. 1–21. DOI: 10.1007/s11229-019-02096-x.

Neta, Ram (2012). "Knowing From the Armchair That Our Intuitions Are Reliable." In: *The Monist* 95.2, pp. 329–351. DOI: 10.5840/monist201295218.

Nichols, Shaun, and Joshua Knobe (2007). "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions." In: *Noûs* 41.4, pp. 663–685. DOI: 10.1111/j.1468-0068.2007.00666.x.

Niiniluoto, Ilkka (2014). "Scientific Progress as Increasing Verisimilitude." In: *Studies in History and Philosophy of Science Part A* 46, pp. 73–77. DOI: 10.1016/j. shpsa.2014.02.002.

Nolan, Daniel (1997). "Quantitative Parsimony." In: *British Journal for the Philosophy of Science* 48.3, pp. 329–343. DOI: 10.1093/bjps/48.3.329.

Novaes, Catarina Dutilh (2018). "Carnapian Explication and Ameliorative Analysis: A Systematic Comparison." In: *Synthese* 197.3, pp. 1011–1034. DOI: 10.1007/ s11229-018-1732-9.

Olsson, Erik J. (2015). "Gettier and the Method of Explication: A 60 Year Old Solution to a 50 Year Old Problem". In: *Philosophical Studies* 172.1, pp. 57–72. doi:10.1007/s11098-014-0383-z.

Pacer, M., and Tania Lombrozo (2017). "Ockham's Razor Cuts to the Root: Simplicity in Causal Explanation." In: *Journal of Experimental Psychology: General* 146.12, pp. 1761–1780. DOI: 10.1037/xge0000318.

Pinder, Mark (2017). "Does Experimental Philosophy Have a Role to Play in Carnapian Explication?" In: *Ratio* 30.4, pp. 443–461. DOI: 10.1111/rati.12164.

– (2020). "What Ought a Fruitful Explicatum to Be?" In: *Erkenntnis*, pp. 1–20. DOI: 10.1007/s10670-020-00223-6.

Plunkett, David, and Herman Cappelen (2020). "A Guided Tour Of Conceptual Engineering and Conceptual Ethics." In: *Conceptual Engineering and Conceptual Ethics*. Ed. by Herman Cappelen, David Plunkett, and Alexis Burgess. Oxford: Oxford University Press.

Popper, Karl (1959). *The Logic of Scientific Discovery*. Routledge.

Prinz, Jesse J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. MIT Press.

Prinzing, Michael (2018). "The Revisionist's Rubric: Conceptual Engineering and the Discontinuity Objection." In: *Inquiry: An Interdisciplinary Journal of Philosophy* 61.8, pp. 854–880. doi: 10.1080/0020174x.2017.1385522.

Quine (1986). "reply to white". In: *in L. Hahn and P. Schilpp (eds.), The Philosophy of W.V. Quine, La Salle, Ill.: Open Court*, pp. 663–665.

Quine, Willard van (1963). "On Simple Theories of a Complex World". In: *Synthese* 15.1, pp. 103–106. doi: 10.1007/BF00484843.

Railton, Peter (2003). *Facts, Values, and Norms: Essays Toward a Morality of Consequence*. Cambridge University Press.

Rawls, John (1980). "Kantian Constructivism in Moral Theory". In: *Journal of Philosophy* 77.9, pp.515-572. doi: 10.2307/2025790.

Saul, Jennifer (2006). "Gender and Race". In: *Aristotelian Society Supplementary Volume* 80.1, pp. 119–143. doi: 10.1111/j.1467-8349.2006.00140.x.

Scharp, Kevin (2013). *Replacing Truth*. Oxford University Press UK.

Schupbach, Jonah N. (2017). "Experimental Explication." In: *Philosophy and Phenomenological Research* 94.3, pp. 672–710. DOI: 10.1111/phpr.12207.

Schwitzgebel, Eric and Fiery Cushman (2015). "Philosophers 'Biased Judgments Persist Despite Training, Expertise, and Reflection." In: *Cognition* 141, pp. 127–137. DOI: 10.1016/j.cognition.2015.04.015.

Seyedsayamdost, Hamid (2015). "On Normativity and Epistemic Intuitions: Failure of Replication." In: *Episteme* 12.1, pp. 95–116. DOI: 10.1017/epi.2014.27.

Shan, Yafeng (2019). "A New Functional Approach to Scientific Progress." In: *Philosophy of Science* 86.4, pp. 739–758. DOI: 10.1086/704980.

Shepherd, Joshua and James Justus (2015). "X - Phi and Carnapian Explication". In: *Erkenntnis* 80.2, pp. 381–402. doi: 10.1007/s10670-014-9648-3.

Sinatra, Gale M. and Paul R. Pintrich (2003). *Intentional Conceptual Change*. L. Erlbaum.

Sinnott-Armstrong, Walter (Jan. 2008). "Framing Moral Intuitions." In: *Moral Psychology* 2.

Sober (2000). *Philosophy of Biology*. Westview Press.

– (2010). "Evolutionary Theory and the Reality of Macro Probabilities." In: *The Place of Probability in Science*. Ed. by Ellery Eells and James H. Fetzer. Springer, pp. 133–60.

– (2015). *Ockham's Razors: A User's Manual*. Cambridge University Press.

– (1990). "Explanation in Biology: Let's Razor Ockham's Razor." In: *Royal Institute of Philosophy Lectures* 27, pp. 73–93.

Strawson, Peter F. (1963). "Carnap's Views on Conceptual Systems Versus Natural Languages in Analytic Philosophy." In: *The Philosophy of Rudolf Carnap*. Ed. by Paul Arthur Schilpp. Open Court: La Salle, pp. 503–518.

Swain, Stacey, Joshua Alexander, and Jonathan M. Weinberg (2008). "The Instability of Philosophical Intuitions: Running Hot and Cold on Truetemp." In: *Philosophy and Phenomenological Research* 76.1, pp. 138–155. DOI: 10.1111/j.1933-1592. 2007.00118.x.

Thagard, Paul (1992). *Conceptual Revolutions*. Princeton University Press.

Thomasson, Amie (Jan. 2020). "A Pragmatic Method for Normative Conceptual Work." In: *Conceptual Engineering and Conceptual Ethics,* pp. 435–458.

Thornhill, Randy and Steven Gangestad (2008). *The evolutionary biology of human female sexuality*. New York: Oxford University Press.

Tobia, Kevin, Wesley Buckwalter, and Stephen Stich (2013). "Moral Intuitions: Are Philosophers Experts?" In: *Philosophical Psychology* 26.5, pp. 629–638. DOI: 10.1080/09515089.2012.696327.

Weinberg, Jonathan M. (2007). "How to Challenge Intuitions Empirically Without Risking Skepticism." In: *Midwest Studies in Philosophy* 31.1, pp. 318–343. DOI: 10.1111/j.1475-4975.2007.00157.x.

Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich (2001). "Normativity and Epistemic Intuitions." In: *Philosophical Topics,* 29.1-2, pp. 429–460. DOI: 10.5840/philtopics2001291/217.